HHS Public Access

Author manuscript

Cell. Author manuscript; available in PMC 2020 October 05.

Published in final edited form as:

Cell. 2019 October 17; 179(3): 713–728.e17. doi:10.1016/j.cell.2019.09.020.

Multimodal Analysis of Cell Types in a Hypothalamic Node Controlling Social Behavior

Dong-Wook Kim[1,2,4], **Zizhen Yao**[5], **Lucas T. Graybuck**[5], **Tae Kyung Kim**[5], **Thuc Nghi Nguyen**[5], **Kimberly A. Smith**[5], **Olivia Fong**[5], **Lynn Yi**[2], **Noushin Koulena**[2], **Nico Pierson**[2], **Sheel Shah**[2], **Liching Lo**[2,3,4], **Allan-Hermann Pool**[2], **Yuki Oka**[2], **Lior Pachter**[2], **Long Cai**[2], **Bosiljka Tasic**[5], **Hongkui Zeng**[5], **David J. Anderson**[2,3,4,6,*]

[1]Program in Computation and Neural Systems, California Institute of Technology, Pasadena, CA, USA

[2]Division of Biology and Biological Engineering 156-29, California Institute of Technology, Pasadena, CA, USA

[3]Howard Hughes Medical Institute

[4]Tianqiao and Chrissy Chen Institute for Neuroscience, California Institute of Technology, Pasadena, CA 91125, USA

[5]Allen Institute for Brain Science, Seattle, WA 98109, USA

[6]Lead contact

SUMMARY

The ventrolateral subdivision of the ventromedial hypothalamus (VMHvl) contains ~4,000 neurons that project to multiple targets and control innate social behaviors including aggression and mounting. However, the number of cell types in VMHvl and their relationship to connectivity and behavioral function are unknown. We performed single-cell RNA sequencing using two independent platforms—SMART-seq (~4,500 neurons) and 10x (~78,000 neurons)—and investigated correspondence between transcriptomic identity and axonal projections or behavioral activation, respectively. Canonical correlation analysis (CCA) identified 17 transcriptomic types (T-types), including several sexually dimorphic clusters, the majority of which were validated by seqFISH. Immediate early gene analysis identified T-types exhibiting preferential responses to intruder males versus females but only rare examples of behavior-specific activation.

*Correspondence: wuwei@caltech.edu.

AUTHOR CONTRIBUTIONS

D.-W.K., B.T., H.Z., and D.J.A. contributed to the study design. D.-W.K. performed most of the experiments. T.K.K. and K.A.S. prepared sequencing libraries for SMART-seq scRNA-seq. L.T.G and O.F. contributed data visualization. D.-W.K. and T.N.N performed Retro-seq experiments. D.-W.K., Z.Y, L.T.G, L.Y., and L.P. analyzed the scRNA-seq data. D.-W.K. and N.K. performed seqFISH experiments. D.-W.K., N.P., S.S., and L.C. analyzed the seqFISH data. D.-W.K. and L.L. performed retrograde labeling with *c-fos* immunohistochemistry experiments. A.-H.P. and Y.O. developed the tissue preparation protocols for 10x Act-seq. D.J.A. supervised the project. D.-W.K. and D.J.A. wrote the manuscript with contributions from B.T. and H.Z. All authors discussed and commented on the manuscript.

SUPPLEMENTAL INFORMATION

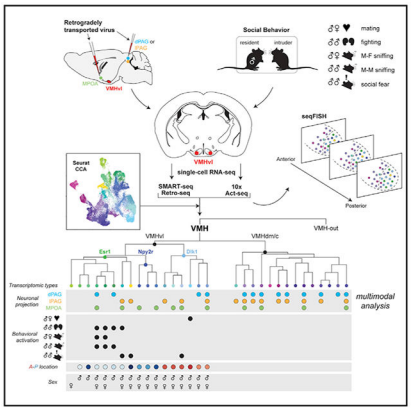
Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2019.09.020.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Unexpectedly, many VMHvl T-types comprise a mixed population of neurons with different projection target preferences. Overall our analysis revealed that, surprisingly, few VMHvl T-types exhibit a clear correspondence with behavior-specific activation and connectivity.

Graphical Abstract



In Brief

Single-cell RNA sequencing of the mouse ventrolateral subdivision of the ventromedial hypothalamus, a brain region that contains ~4,000 neurons and controls innate social behaviors including aggression and mounting, reveals several transcriptomic cell types that differ between males and females.

INTRODUCTION

Cell types are a fundamental unit of organization and specificity in multicellular organisms. An understanding of cellular diversity in the brain is critical for studies of neural function and dysfunction (Jorgenson et al., 2015). Using single-cell RNA sequencing (scRNA-seq), recent surveys have estimated up to ~600 different transcriptomic cell types (T-types) in the mouse brain (Saunders et al., 2018; Zeisel et al., 2018). Such diversity immediately poses the “correspondence problem:” how is transcriptomic heterogeneity related to other facets of neuronal identity, such as connectivity and physiology (Tasic, 2018; Zeng and Sanes, 2017)? While such correspondence is well established for retinal cell types (Macosko et al., 2015; Seung and Sümbül, 2014), it is not yet clear whether this principle extends to the central brain. For example, out of over a hundred cortical T-types (Tasic et al., 2018), only two have been shown to project to different subcortical targets and to exert distinct functions in motor control (Economo et al., 2018).

The hypothalamus is an evolutionarily ancient collection of deep subcortical nuclei that control homeostatic and innate “survival behaviors” and associated motivational states (reviewed in Luiten et al., 1987; Saper and Lowell, 2014; Sternson, 2013). The instinctive nature of these functions suggests they might be controlled by specific T-types with genetically specified connectivity. Initial scRNA-seq studies have revealed evidence of extensive hypothalamic cell diversity (2–35 cell types/mm^3 tissue sampled) (Campbell et al.,

2017; Chen et al., 2017; Romanov et al., 2017), but the behavioral relevance of such cell types was not examined. A recent study of the preoptic region (POR), a large area (~20% of total hypothalamus volume) containing ~20 distinct subdivisions (nuclei), revealed ~70 T-types (~3–4 T-types per nucleus). Multiplexed error-robust fluorescence *in situ* hybridization (MERFISH) (Moffitt et al., 2016) experiments using the immediate early gene (IEG) *c-fos* (Greenberg and Ziff, 1984; Morgan et al., 1987) indicated that some T-types were preferentially activated during a particular social behavior (Moffitt et al., 2018). However, the relationship between transcriptomic identity and axonal projections (Kohl et al., 2018) was not investigated.

The ventromedial hypothalamus of the ventrolateral subdivision (VMHvl) occupies ~0.5% of hypothalamus volume and contains ~4,000 primarily glutamatergic neurons that collectively control social behaviors, including aggression, as well as metabolism (reviewed in Chen and Hong, 2018; Hashikawa et al., 2017b; Kennedy et al., 2014; Krause and Ingraham, 2017). Calcium imaging of estrogen receptor type 1-expressing VMHvl (VMHvl^{Esr1}) neurons during social behavior revealed that population activity represents intruder sex identity and behavior (Remedios et al., 2017), while anatomic analysis indicated that VMHvl^{Esr1} neurons project to multiple (~30) downstream targets (Lo et al., 2019). Cellular subpopulations in VMHvl have been identified using morphology (Millhouse, 1973a; b), molecular markers (Correa et al., 2015; Xu et al., 2012) , and bulk RNA-seq (Hashikawa et al., 2017a), but the behavioral function and connectivity of these populations were not established.

We have carried out scRNA-seq of VMHvl neurons at high sampling density (2×10^5 neurons sequenced/mm^3 of tissue), using two independent platforms: SMART-seq v4 (Picelli et al., 2013) and the 10x genomics droplet-based platform (Zheng et al., 2017). Our results identify 17 T-types in this subnucleus alone, a density (~160 T-types/mm^3 tissue sampled) 3- to 68-fold higher than reported in previous studies (Chen et al., 2017; Mickelsen et al., 2019; Moffitt et al., 2018; Romanov et al., 2017). Investigation of correspondence between anatomy, behavioral activation, and projection specificity using seqFISH (Eng et al., 2019; Shah et al., 2016), Act-seq (Wu et al., 2017), and Retro-seq (Tasic et al., 2018) indicated that, with a few notable exceptions, most VMHvl T-types do not map to specific behaviors and project to specific targets.

RESULTS

A Census of VMHvl Transcriptomic Cell Types Using SMART-Seq

Initially, we used SMART-seq (v4, Clontech) to derive a census for VMHvl cell types. Because of its small size (Figure S1A) and the difficulty of manually dissecting it free from surrounding tissue, we pre-selected for neurons in this nucleus using transgenic Cre lines, fluorescent Cre-dependent reporters, and fluorescence activated cell sorting (FACS; Figure S1B) (Tasic et al., 2016; Tasic et al., 2018). This approach yielded ~100 cells per animal on average. We used several Cre lines (see STAR Methods; Figures S1C–S1E); *Nr5a1-Cre* was particularly useful because it afforded labeling of ~75% of all cells in adult VMHvl (Figures S1G and S1H; *Nr5a1* is transiently expressed in most VMH cells during development but is downregulated in adult VMHvl [Cheung et al., 2013]). We collected 4,037 cells from 48

males and 537 cells from six females (Figure S1C) under home cage or a variety of behavioral conditions (Figures S1D and S1E). In total, 4,574 cells were sequenced from 54 mice, of which 4,473 were neurons (Figure 1C).

Clustering using the released version of scratch.hicat (Tasic et al., 2018) yielded 46 transcriptomic "leaf" clusters, or transcriptomic cell types (T-types), at the terminal hierarchical level (Figure 1A), six of them non-neuronal and the remainder neuronal (Figure 1C). A similar number of clusters was identified using two other independent methods: Ward's hierarchical clustering (Ward, 1963) or graph-based clustering using Seurat (Butler et al., 2018) (Figures S2D and S2E). We used differentially expressed genes (DEGs) and prior knowledge to group neuronal T-types into several broad categories with different predicted anatomic locations. VMH has a core-and-shell organization, such that cells within the core are mostly glutamatergic and express *Fezf1*, *Adcyap1*, *Gda*, *Nrgn*, *Cbln1*, *Lmo3*, *Nr5a1*, *C1ql2*, and *Rreb1* (Kurrasch et al., 2007), while those in the surrounding shell are mostly GABAergic (Choi et al., 2005; Hahn et al., 2019) and express *Gad1*, *Gad2*, *Tmem176a*, *Tmem176b*, *Six3*, *Dlk1*, and *Ecel1* (Figure 1B, upper). Only 13 of 40 neuronal T-types (n = 611 out of 4,473) expressed *Gad1*, *Gad2*, and *Slc32a1 (Vgat)*, reflecting the fact that our Cre driver lines were chosen to label cells within VMH (Figures 1C, S1F, and S1G). Within VMH, two sub-populations could be defined: those in VMHc or anterior VMH, which expressed *Nr5a1*, *C1ql2*, *Rreb1*, *Six3*, and *Ldb2*; and those in VMHvl (Figure 1B, lower). There were few T-types from dorso-medial VMH (VMHdm), as our microdissection procedure deliberately selected against this region (Figure S1B, yellow outline). Within VMHvl, several classes of neurons could be identified, by expression of *Esr1*, *Dlk1* (Deltalike homolog 1; also called Pre-adipocyte factor 1, or *Pref-1*), and *Satb2* (Figures 1B, S2A, and S2C). Altogether, this analysis identified 17 different T-types predicted to be in VMHvl.

Anatomic Analysis of T-Types by seqFISH

To place the T-types identified by SMART-seq in anatomical context with higher resolution, we performed sequential fluorescence *in situ* hybridization (seqFISH) (Eng et al., 2019; Shah et al., 2016). We applied 58 probes comprising DEGs that marked the different VMHvl SMART-seq clusters and carried out multiple sequential cycles of single-molecule (sm) FISH on sections covering multiple positions along the anterior-posterior axis (Figures 2A1 and 2G). Over 23,000 neurons were semi-automatically segmented within four regions of interest (ROIs) per section (Figure 2A2), and hybridization signals for each probe set were counted within each cell. Segmented cells were distributed across VMHvl, VMHdm/c, and surrounding regions (VMH-out; Figures 2A3 and S3).

We performed independent clustering of the seqFISH data (Figure 2B) and compared the results to the SMART-seq cluster analysis. In general, there was a high correlation between VMHvl SMART-seq clusters and seqFISH clusters (Figure 2C; 85.2% of seqFISH clusters correspond to at least one SMART-seq T-type), as well as of marker gene expression levels between the two datasets (Figures S3F and S3G), suggesting that transcriptomic profiles were not strongly altered by tissue dissociation or FACS. In general, the anatomical locations of cells revealed by seqFISH validated the provisional locations of SMART-seq

clusters assigned by marker expression Figures 1 and S3). However, we found several cases where cells in certain clusters "violated" cytoarchitectonic locations predicted by marker expression: (1) a small number of GABAergic cells assigned to VMH-out clusters by SMART-seq was detected just inside the lateral border of VMHvl (Figures S3A [VMH-out], S3B, and S3C), (2) a small number of cells assigned to VMHdm/c (based on expression of *Nr5a1* and *C1ql2*) were found just across the border in VMHvl (Figures S3A [VMHdm/c], S3D, and S3E), (3) some cells assigned to VMHvl were also detected in the adjacent VMH-out region (Figures S3A [VMHvl], S3D, and S3E). Interestingly, cells in several of the VMHvl seqFISH clusters exhibited different distributions along the anterior-posterior axis, e.g., those in seqFISH cluster #11 (Esr1_2,3) were located more posteriorly, while those in cluster #5 (Nr5a1_4,6) were located more anteriorly (Figure 2D, lower).

Neuronal Projections of VMHvl T-Types

Classical neuroanatomical studies have indicated that VMHvl neurons, as a population, project to multiple (>20) brain regions (Canteras et al., 1994; Saper et al., 1976). The relationship between these projections and different VMHvl cell types is not yet clear. Recent studies using viral tracing in *Esr1-Cre* mice have identified two anatomically distinct subsets of VMHvl^{Esr1} neurons that collateralize preferentially to posterior versus anterior targets, such as the dorso-medial periaqueductal gray (dmPAG) versus the medial preoptic area (MPOA), respectively (Lo et al., 2019). To examine the relationship of VMHvl T-types to these projection-specific subsets, we performed scRNA-seq following retrograde labeling from MPOA or dorsal or lateral peri-aqueductal gray (dPAG and lPAG, respectively) (Retro-seq; see STAR Methods; Figure 3A) (Tasic et al., 2018).

These data revealed groups of VMH T-types with clear projection preferences. Consistent with anterograde labeling studies (Kunwar et al., 2015; Oh et al., 2014; Wang et al., 2015), the *Nr5a1*$^{+}$ clusters, which are located primarily in VMHdm/c, showed a significant projection bias to the dPAG (Figure 3B). By contrast, the *Esr1*$^{+}$ and *Dlk1*$^{+}$ classes, which are located in VMHvl, showed a significant projection bias to the MPOA. This is not surprising, as MPOA is one of the strongest projection targets of VMHvl (Canteras et al., 1994; Lo et al., 2019; Roeling et al., 1994; Yang et al., 2013). Surprisingly, although retrograde labeling in *Esr1-Cre* mice revealed mostly non-overlapping populations of MPOA- or dPAG-projecting VMHvl^{Esr1} neurons (Lo et al., 2019), most *Esr1*$^{+}$ T-types in VMHvl contained cells back labeled from both targets, indicating a probable mixture of cells with different projections. Nevertheless a few T-types exhibited a dPAG projection preference (#28/Esr1_4 and #39/Dlk1_5; Figures 3C and 3D). Cluster #31 (*Nup62cl*$^{+}$) was specifically labeled by injections to lPAG. Some T-types were underrepresented in the retrogradely labeled populations and may correspond to local interneurons or to projection biases not interrogated by our retrograde injection sites (Figure 3D, bars below dashed line).

Linking Transcriptomic Identity to Activity by Act-Seq

Electrophysiology, *c-fos* catFISH *in situ* hybridization (Lin et al., 2011), and *in vivo* calcium imaging (Remedios et al., 2017) studies have revealed that VMHvl in males contains distinct, largely non-overlapping, populations of neurons that are active during social interactions with males versus females. In addition, different patterns of VMHvl neuronal

activity are observed during mating versus aggression (Falkner et al., 2014; Hashikawa et al., 2017a; Remedios et al., 2017), as well as during social fear (Sakurai et al., 2016; Silva et al., 2013; Wang et al., 2019). We therefore wished to investigate how VMHvl activity during different social behaviors is related to T-types. Attempts to measure IEG transcript levels using SMART-seq in cells obtained from animals following different social behaviors were unsuccessful due to the low yield of cells per animal after FACS isolation. While this work was in progress, methods for behavioral IEG analysis using droplet-based scRNA-seq were described (Hrvatin et al., 2018; Wu et al., 2017). In order to apply one such method, Act-seq (Wu et al., 2017), to VMHvl, we first characterized cellular diversity in this region using the 10x Genomics droplet-based platform.

We microdissected VMHvl from acute slices, dissociated it using modified protocols designed to maximize cell viability and number of genes detected (A.H.P. and Y.O., in preparation), and immediately subjected the cell suspension to 10x preparation and sequencing without employing FACS. Out of 149,663 cells sequenced (n = 90 mice), 27.7% were VMH neurons (*Slc17a6*$^{+}$, *Fezf1*$^{+}$, and *Adcyap1*$^{+}$), 24.8% were non-VMH neurons (*Slc32a1*$^{+}$, *Gad1*$^{+}$, and *Gad2*$^{+}$), and 47.5% were non-neuronal cells (Figures S4B–S4I). Higher proportions of non-VMH neurons, compared with the SMART-seq dataset, are expected since FACS enrichment for VMH markers was omitted.

Analysis of the VMH neurons revealed 29 T-types and their DEGs (Figures 4E and S4L). Of these, 17 clusters (including one female-specific cluster; see below) were provisionally assigned based on marker expression to VMHvl (Figure 4E, red squares) and the remainder to VMHdm/c (Figure 4E, blue squares). Like SMART-seq, the 10x data revealed that VMHvl neurons fell into three distinct classes defined by *Esr1*, *Satb2*, and *Dlk1* (Figure 4D).

To determine how neural activation during different behaviors maps onto 10x-derived VMHvl T-types, we repeated the 10x analysis using tissue dissected from animals ~1 h after they engaged in one of eight different social behaviors (see STAR Methods; Figures 5A and S4I). Close investigation (CI) was tested using intruders either suspended by the tail or confined in a mesh pencil cup. Initially, for each cluster and each behavior, we compared the percentage of *Fos*$^{+}$ cells and the average fold-change in *Fos* expression versus control (Figure 5B). This analysis revealed that 8 of 16 male VMHvl T-types (50%), in aggregate, displayed a statistically significant *Fos* induction during the behaviors tested. During most behaviors, one to three T-types were activated (with the exception of social fear in single-housed males; seven T-types), and each of those T-types was typically activated during multiple behaviors. However, a few T-types showed more restricted patterns of activation. For example, cluster Nup62cl^{10x} was activated only during social fear, while cluster Dlk1_1^{10x} was only activated during mating (Figure 5B).

In addition to exhibiting higher *Fos* expression during a particular behavior in comparison with home-cage control animals, some T-types also exhibited significant differences in *Fos* expression between different behaviors (Figure S5A, rows, colored squares). For example, cluster Esr1_7^{10x} was activated to a significantly greater extent during aggression than during male-male (M-M) CI (suspended by tail), social fear (group-housed mice), male-female (M-F) CI, or mating, while cluster Dlk1_1^{10x} was activated at higher levels during

mating than during aggression or M-M CI (pencil cup). Seven T-types were activated during social fear in single-housed mice, but only three were activated during social fear in group-housed mice (Figure 5B). This difference is likely to reflect defensive aggression, which is initially displayed by most single- but not group-housed mice.

Previous studies have shown that distinct, partially overlapping subsets of VMHvl neurons are activated in males during fighting versus mating, using both *Fos* cellular compartment analysis of temporal activity by fluorescence *in situ* hybridization (catFISH) (Lin et al., 2011) and calcium imaging of GCaMP6 expressed in *Esr1-Cre* mice (Remedios et al., 2017). While cluster Esr1_7^{10x} was activated during M-M but not M-F social interactions, none of the other *Esr1*$^{+}$ clusters showed significant *Fos* induction during mating relative to controls (Figure 5B). However, statistical correction for multiple comparisons in this analysis might have obscured the relatively weaker *Fos* induction during mating (Lee et al., 2014) in some *Esr1*$^{+}$ clusters. To investigate this possibility, we performed a statistical contrast between combined M-F social interaction tests (mating and two M-F CI tests), M-M interaction tests (aggression and two M-M CI tests), and controls among the different *Esr1*$^{+}$ clusters. Because Act-seq allows simultaneous analysis of multiple IEGs (Wu et al., 2017), we examined *Fos*, *Fosl2*, and *Junb* expression. This analysis indicated that cluster Esr1_5^{10x} exhibited significantly greater *Fosl2* induction than cluster Esr1_7^{10x} following M-F social interactions (Figure 5C, *Fosl2*, blue data points), while conversely Esr1_7^{10x} showed significantly greater *Fosl2* induction than cluster Esr1_5^{10x} following M-M social interactions (Figure 5C, red data points). Cluster Esr1_4^{10x} was activated during both types of interactions. Similar trends were observed for *Fos* and *Junb* expression.

Systematic examination of 139 different IEGs (Bravo, 1990; Wu et al., 2017) revealed different patterns of IEG activation in different clusters during different behaviors (Figures 5D and S5B). Cluster Esr1_7^{10x} exhibited activation of multiple IEGs during aggression and M-M CI tests, cluster Dlk1_1^{10x} showed the most IEGs induced during mating, and cluster Esr1_4^{10x} was activated to varying extents during all eight behaviors (Figure 5D).

We investigated next whether VMHvl T-types activated during social behaviors were also distinguished by other classes of non-activity-dependent DEGs. Interestingly, certain classes of DEGs were almost as effective, as the entire set of annotated mRNAs at classifying these behaviorally relevant clusters (Figure S6A), including the I-set domain-containing group of immunoglobulin (Ig) superfamily cell surface proteins (Figure S6B, mean area under receiver-operator characteristic [AUROC] curve = 0.92 ± 0.05), while others including neuropeptides, Class A orphan G protein-coupled receptors (GPCRs), and transcription factors were less discriminating (Table S3).

Alignment of SMART-Seq and 10x Datasets by CCA

Next, we sought to determine whether there was a three-way correspondence between T-type identity, behavioral activation. and projection specificity. As the latter two measures were carried out using different scRNA-seq platforms, we first aligned the SMART-seq and 10x datasets by canonical correlation analysis (CCA; Figure S7) (Butler et al., 2018; Stuart et al., 2019). Co-clustering of the joint dataset revealed 31 CCA VMH-in clusters (Figure S7C).

Both the 10x and SMART-seq clusters were relatively well-aligned with the joint CCA clusters, with some exceptions (Figures S7D, S7E, and S7G).

We then computed the hierarchical relationship between and compared the expression of different marker genes among these joint CCA clusters (Figure 6A). 17 glutamatergic CCA clusters were predicted to be in VMHvl and the remainder in VMHdm/c, based on the expression of *Fezf1*, *Adcyap1*, *Slc17a6*, and *Nr5a1* (see above). We found excellent agreement between the cluster markers identified by 10x versus SMART-seq (Figure 6A). VMHvl neurons again fell into 3 distinct subsets marked by *Esr1*, *Satb2*, and *Dlk1* (Figures 6A and S7H).

Identification of Sexually Dimorphic CCA Clusters in VMHvl

We annotated seven joint CCA clusters as *"Esr1$^+$"* (Figures 6A and S7H). One of these seven clusters, which expressed higher levels of *Esr1* than the others, was female specific (Tsix_Esr1^{CCA3} Figures 6B, 6C [this cluster was split into two subclusters by SMART-seq], and S7G). Conversely, clusters Esr1_1^{CCA4} and Esr1_2/3^{CCA5} (Figures 6B and 6C) were strongly enriched in males; the latter expressed *Moxd1* (Figure 6B), which encodes an enzyme expressed dimorphically in the male hypothalamus (Tsuneoka et al., 2017). To validate these results, we performed seqFISH analysis for additional markers identified by scRNA-seq in males and females using *Esr1/Gpc3* or *Gldn* as reference markers (Figures 6B and 6F). The results confirmed the differential expression of these additional markers in these sexually dimorphic cell types and indicated a posterior bias in the anatomic location of their cell somata (Figures 6E and 6F).

A previous study demonstrated that female VMHvl contains two distinct subpopulations of *Esr1*$^+$ neurons: one, activated during mating, is located in a subregion called VHMpvll; the other, activated during aggression, is located in subregion VMHpvlm (Hashikawa et al., 2017a). Marker genes for these two anatomic subdivisions were identified using bulk RNA-seq (Hashikawa et al., 2017a). Comparison of those marker genes with our scRNA-seq dataset revealed that the female-specific cluster Tsix_Esr1^{CCA3} expressed most of the markers spatially enriched in VMHpvll (mating region) (Figure 6D), while cells in the *Nr5a1*$^+$ and *Dlk1*$^+$ classes expressed most of the markers enriched in VMHpvlm (aggression region).

Relationship between Behavioral Activation, Projection Specificity, and Transcriptomic Identity

We next sought to correlate transcriptomic phenotype, behavioral function, and projection bias (Economo et al., 2018; Tasic et al., 2018). Although it was necessary to measure projection bias and behavioral activation using two different scRNA-seq platforms (SMART-seq and 10x, respectively), the correspondence provided by CCA (Stuart et al., 2019) allowed us to correlate the two phenotypic measures. Surprisingly, we did not observe a clear 1-to-1-to-1 relationship between T-type identity, behavioral activation, and projection bias for most clusters (Figures 7A and 7E). A rare exception was T-type Nup62cl, which was selectively activated during social fear (in group-housed mice), and which selectively projected to lPAG (Figure 7A, yellow), a known circuit node controlling freezing behavior in

mice (Tovote et al., 2016). Similarly, another social fear-activated T-type, Dlk1_4^{CCA18}, also preferentially projected to lPAG (Figure 7A). In general, there was a positive correlation (r = 0.66) between the proportions of cells projecting to lPAG and stronger *Fos* expression induced by social fear (Figures 7C and 7E).

We also correlated preferential activation during M-M versus M-F social interactions with projection bias. A positive correlation was observed between CCA T-types that were preferentially activated during all M-M versus all M-F social interactions (e.g., Esr1_7^{CCA9}; Figure 7A, all M-M), and that exhibited a relative bias for projections to dPAG versus MPOA based on Retro-seq data (Figure 7B; r = 0.71). By contrast, cluster Esr1_5^{CCA7}, which exhibited significantly greater *Fosl2* inductions following M-F versus M-M social interactions (Figure 7A), did not show any projection bias.

To independently investigate the relationship between social behavior and projection bias for VMHvl^{Esr1} neurons, we performed experiments combining retrograde tracing from dPAG versus MPOA with anti-Fos immunostaining of *Esr1*$^{+}$ neurons following aggression, mating, or M-F CI (see STAR Methods). The results of these experiments indicated that VMHvl^{Esr1} cells activated during aggression showed a significant bias for projection to dPAG versus MPOA, whereas VMHvl^{Esr1} cells activated following M-F interactions showed no such bias (Figure 7D).

These latter observations suggested that we should find a CCA cluster preferentially activated during aggression, which contained only dPAG-projecting cells; however, we did not. Cluster Esr1_7^{CCA9} showed strong, specific IEG activation during aggression (Figures 5D and 7A) but corresponded to a SMART-seq cluster (Esr1_4 Smoc2^{SS}) that contains dPAG- as well as MPOA-projecting cells by Retro-seq (Figure 3C). However, Esr1_4 Smoc2^{SS} was mapped by CCA not only to Esr1_7^{CCA9} but also to Esr1_4^{CCA6} (Figure S7G); the latter cluster did not show any preferential activation during aggression versus other social behaviors (Figures 5D and 7A). If the dPAG-projecting and MPOA-projecting cells in cluster Esr1_4^{SS} are indeed distinct (Lo et al., 2019), then the dPAG projection-specific cells in that cluster could correspond to those in the aggression-activated cluster Esr1_7^{CCA9}. If so, then Esr1_7^{CCA9} might indeed represent an aggression-activated T-type that preferentially projects to dPAG and other brainstem structures (Lo et al., 2019). Notably, this cluster was slightly enriched in males versus females (Figure 6B, #9).

DISCUSSION

The hypothalamus contains ~65 nuclei or regions per hemisphere (Hahn et al., 2019). Earlier scRNA-seq studies identified ~30–60 T-types in the entire hypothalamus (Chen et al., 2017; Romanov et al., 2017). Recently, Moffitt et al. (2018) reported ~70 cell types in the POR, which contains ~20 nuclei (a density of 3–4 cell types per nucleus). Our results reveal ~30 T-types in VMH and 17 T-types in VMHvl alone. Similarly, ~35 T-types were identified in the arcuate nucleus, which controls feeding and is of comparable size to VMH (Campbell et al., 2017). These results suggest that the total number of hypothalamic cell types has been significantly underestimated.

Our results demonstrate good agreement between T-types identified by the 10x and SMART-seq platforms. We identified three major transcriptomic classes in VMHvl. One is defined by *Esr1* (Morrell and Pfaff, 1982), the other two by *Dlk1/Pref-1* (Persson-Augner et al., 2014) and *Satb2,* respectively. The *Esr1*$^{+}$ class, which expresses the progesterone receptor (*Pgr*) and controls social behaviors (Lee et al., 2014; Yang et al., 2013; Yang et al., 2017), contains seven different T-types, while the *Dlk1*$^{+}$ class contains six. Three *Satb2* T-types and Nup62cl comprise the remaining clusters. While we identified some VMHdm/c T-types, which contaminated our VMHvl dissection and which include *Sf1/Nr5a1*$^{+}$ neurons, a comprehensive census of this regions remains to be performed.

Using a combination of previously defined and newly discovered markers, we performed seqFISH on VMHvl *in vivo*. Independent clustering of the seqFISH data confirmed the cluster diversity revealed by SMART-seq and 10x–only 4 of 27 seqFISH clusters (15%) did not map to any scRNA-seq T-types. Several of the seqFISH clusters exhibited differences in cellular distribution along the A-P axis of VMHvl, consistent with earlier observations (Lo et al., 2019; Wang et al., 2019). Our results are also suggestive of heterogeneity along the M-L axis of VMHvl (Hashikawa et al., 2017a). Thus, some T-types appear spatially restricted within VMHvl.

Male- and Female-Enriched *Esr1*$^{+}$ Neuronal Cell Clusters in VMHvl

Sex-specific neurons are well established in *Drosophila* (Cachero et al., 2010; Yu et al., 2010) and *C. elegans* (Liu and Sternberg, 1995), but have not previously been reported in mice. While quantitative sex differences in gene expression have been described in rodents (Xu et al., 2012), including at the single-cell level (Welch et al., 2019), our results provide initial evidence of sex-specific neuronal T-types in the mammalian CNS. Importantly, these sex-specific T-types are not defined simply by sex-chromosome-encoded sexually dimorphic genes (e.g., *Tsx1*, which is also expressed in most non-sexually dimorphic T-types in females). Rather, they are defined by specific patterns of covariation of autosomal gene expression. scRNA-seq analysis at high sampling density is likely to reveal more examples of such rare, sex-specific populations in other sexually dimorphic brain regions (Yang and Shah, 2014).

Foundational studies in the rat have shown that VMHvl^{Esr1} neurons control sexual receptivity (lordosis) in females (Pfaff, 2017; Pfaff and Sakuma, 1979a; b) and were the first to implicate this region and genetic marker in a social behavior. Comparison of our data with bulk RNA-seq data (Hashikawa et al., 2017a) and preliminary Act-seq experiments suggest that the female-specific cluster Tsix_Esr1 may be specifically activated during mating. Functional studies will be required to confirm whether this cell type indeed controls lordosis behavior.

Relationship of Transcriptomic Identity to Projection Heterogeneity in VMHvl

Our results indicate a clear correspondence with projection bias for groups of T-types in VMHdm/c versus those in VMHvl–the former preferentially project to dPAG and lPAG, while the latter project preferentially to the MPOA. This may reflect the fact that VMHvl controls social behaviors, many of which are also controlled by neurons in MPOA

(McHenry et al., 2017; Sano et al., 2013; Simerly, 2002; Wei et al., 2018; Wu et al., 2014). In contrast, VMHdm/c mainly controls anti-predator (non-social) defensive behaviors (Kunwar et al., 2015; Silva et al., 2013; Wang et al., 2015), although one subpopulation in this region controls female sexual receptivity (Ishii et al., 2017).

The correspondence between projection bias and T-types within VMHvl is less clear. We recently identified two non-overlapping subpopulations of VMHvl^{Esr1} neurons which exhibit a strong projection bias to posterior (e.g., dPAG) versus anterior (e.g., MPOA) targets, respectively (Lo et al., 2019). Cells from these two subpopulations also have different cell body sizes and are differentially distributed along the A-P axis. Surprisingly, however, most *Esr1*$^{+}$ T-types contained neurons retrogradely labeled from both dPAG and MPOA, indicating that a single T-type may contain multiple projection-specific types. This may be explained by the transient developmental expression of genes that determine projection specificity as seen in the *Drosophila* antennal lobe (Li et al., 2017). Alternatively, subtle differences in adult gene expression between projection-specific cells within a given T-type may not be detectable by the methods used here.

Functional Heterogeneity among VMHvl Neurons during Social Behaviors

VMHvl neurons have been functionally implicated in the control of aggression, mating, social fear and the encoding of conspecific sex (reviewed in Hashikawa et al., 2017b; Kennedy et al., 2014; Yang and Shah, 2014). In particular, optogenetic activation of VMHvl^{Esr1} neurons can promote mating or aggressive behavior in both males (Lee et al., 2014) and females (Hashikawa et al., 2017a). Our Act-seq data indicated that both distinct and common *Esr1*$^{+}$ T-types were activated during aggression or mating. The common types may control behavioral elements shared by mating and aggression (e.g., CI, chasing). However, calcium imaging of VMHvl^{Esr1} neurons indicated that different populations are activated during male versus female CI; indeed most (50%–60%) of the variance in population activity is explained by intruder sex, with only 10% explained by behavior (Remedios et al., 2017). Consistent with this, distinct T-types were preferentially activated following M-M versus M-F social interactions rather than following specific behaviors. Nevertheless, we identified one *Esr1*$^{+}$ T-type that was more strongly activated during aggression than during other behaviors. We also identified a *Dlk*$^{+}$ T-type specifically activated during mating and several T-types specifically activated during social fear (Sakurai et al., 2016; Silva et al., 2013; Wang et al., 2019).

A recent study of *Fos* activation in the POR by MERFISH concluded that “genetically encoded circuits comprised of transcriptionally distinct neuronal cell types control specific hypothalamic functions” (Moffitt et al., 2018). The apparent difference in conclusions between that study and the present one may reflect the higher diversity of T-types in VMHvl and/or methodological differences. Alternatively, it is possible that the relationship between cell types and behavior encoding is different for nuclei in the POR, which are largely GABAergic (Moffitt et al., 2018), and VMHvl, which is primarily glutamatergic. Answering this question will require correlating transcriptomic identity with activity on a millisecond time-scale, e.g., using calcium imaging followed by gene-specific labeling (Lovett-Barron etal., 2017).

Importantly, VMHvl neurons (including the *Esr1*$^{+}$ subset) also control metabolic function (Musatov et al., 2007; Xu et al., 2011); reviewed in Krause and Ingraham, 2017), and there is evidence that *Esr1*$^{+}$ cell populations controlling reproductive behavior and metabolism can be genetically distinguished in females (Correa et al., 2015). The relationship between the VMHvl transcriptomic cell types described here and metabolic function remains to be explored, although cells in the *Dlk1*$^{+}$ class are good candidates, given data implicating this gene in obesity in mice and humans (Moon et al., 2002; Wermter et al., 2008).

Conclusion

The data presented here constitute one of the few attempts to link transcriptomic diversity in the CNS to projection specificity and behavioral or physiological function (see also Economo et al., 2018). With a few rare exceptions, we generally do not observe a clear, 1-to-1-to-1 correspondence between T-types, projection specificity, and behavioral function. It is possible that most VMHvl T-types reflect facets of cell identity (or cell state) distinct from projection specificity and behavioral function. Alternatively, identifying correspondences between T-types, connectivity, and behavioral activation may require methods with higher spatial and temporal resolution than those employed here. Application of such methods should be facilitated by more specific genetic access to VMHvl subtypes using intersectional transgenic models based on the expression of markers identified in this study.

STAR★METHODS

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for reagents and resources should be directed to and will be fulfilled by the Lead Contact, David J. Anderson (wuwei@caltech.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All procedures were performed in accordance with NIH guidelines and approved by the Institutional Animal Care and Use Committee (IACUC) at the California Institute of Technology (Caltech) and Allen Institute for Brain Science (AIBS). Animals were housed and maintained on a reverse 12 h light-dark cycle with food and water *ad libitum*. For SMART-seq, the following transgenic and reporter lines were used: *Nr5a1*$^{Cre/+}$ (Dhillon et al., 2006), *Oxtr*$^{Cre/+}$ (Daigle et al., 2018), *Vglut2*$^{Cre/+}$ (Jackson Laboratory, stock no. 016963), *Esr1*$^{Cre/+}$ (Lee et al., 2014), *Ai14*$^{Tg/+}$ (Madisen et al., 2010), and *Ai110*$^{Tg/+}$ (Daigle et al., 2018), and maintained on the C57BL/6 background. For 10x Act-seq and seqFISH, wild-type (WT) C57BL/6N male mice (experimental), C57BL/6N female mice (for sexual experience), and BALB/c male mice (intruders) were obtained from Charles River (Burlington, MA). We used 54 mice to collect 4,574 cells for SMART-seq, 90 mice to collect 149,663 cells for 10x (Figures S1C–S1E and S4I), and 7 mice for seqFISH (3 mice with 58 probes and 4 mice with 24 probes, respectively). Animals were euthanized at P50–P110 in this study. Behavior was tested during the dark cycle.

METHOD DETAILS

Single-cell isolation, cDNA amplification and library construction for SMART-seq—We isolated fluorescently labeled single cells from the mouse brain as previously described (Tasic et al., 2016; Tasic et al., 2018). Briefly, adult (8–12-week-old) male mice were anaesthetized with isoflurane and perfused with cold artificial cerebrospinal fluid (ACSF) containing $CaCl_2$ (0.5 mM), glucose (25 mM), HCl (96 mM), HEPES (20 mM), $MgSO_4$ (10 mM), NaH_2PO_4 (1.25 mM), myoinositol (3 mM), N-acetylcysteine (12 mM), N-methyl-d-glucamine (NMDG; 96 mM), KCl (2.5 mM), $NaHCO_3$ (25 mM), sodium l-ascorbate (5 mM), sodium pyruvate (3 mM), taurine (0.01 mM), thiourea (2 mM), bubbled with carbogen gas (95% O_2 and 5% CO_2). The brain was sectioned at 350 μm using vibratome (VT1000S, Leica Microsystems) on ice, and the regions of interest (Figure S1B) were microdissected under a fluorescence dissecting microscope from two consecutive sections (−1.22 to −1.94 from Bregma (Franklin and Paxinos, 2008)). For enzymatic digestion, the dissected tissues were transferred to a microcentrifuge tube and treated with 1 mg/mL pronase (Sigma, P6911-1G) in carbogen-bubbled ACSF for ~70 min at room temperature without mixing in a closed tube. After incubation, the pronase solution was exchanged with cold ACSF containing 1% fetal bovine serum (FBS), The tissue pieces were dissociated into single cells by gentle trituration through Pasteur pipettes with polished tip openings of 600-μm, 300-μm and 150-μm diameter. Single cells were sorted by FACS into 8-well PCR strips containing lysis buffer from SMART-Seq v4 kit with RNase inhibitor (0.17 U/μl), immediately frozen on dry ice, and stored at −80°C. We used the SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (Takara, 634894) to reverse transcribe poly(A) RNA and amplify full-length cDNA according to the manufacturer's instructions. We performed reverse transcription and cDNA amplification for 18 PCR cycles in 8-well strips, in sets of 12–24 strips at a time. At least 1 control strip was used per amplification set, which contained 4 wells without cells and 4 wells with 10 pg control RNA. Control RNA was either Mouse Whole Brain Total RNA (Zyagen, MR-201) or control RNA provided in the SMART-Seq v4 kit. All samples proceeded through Nextera XT DNA Library Preparation (Illumina, FC-131-1096) using Nextera XT Index Kit V2 Set A (Illumina, FC-131-2001). Nextera XT DNA Library prep was performed according to manufacturer's instructions except that the volumes of all reagents including cDNA input were decreased to 0.4 × or 0.5 × by volume.

SMART-seq sequencing data processing and quality control—Libraries were sequenced on an Illumina HiSeq2500 platform (paired-end with read lengths of 50 nt) and Illumina sequencing reads were aligned to GRCm38 (mm10) using a RefSeq annotation gff file retrieved from NCBI on 18 January 2016 (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/all/). Sequence alignment was performed using STAR v2.5.3 (Dobin et al., 2013) in twopass Mode. PCR duplicates were masked and removed using STAR option 'bamRemoveDuplicates'. Only uniquely aligned reads were used for gene quantification. Gene counts were computed using the R GenomicAlignments package (Lawrence et al., 2013) summarizeOverlaps function using 'IntersectionNotEmpty' mode for exonic and intronic regions separately. For the SMART-seq dataset, we only used exonic regions for gene quantification. Cells that met any one of the following criteria were removed: <

100,000 total reads, < 1,000 detected genes (CPM > 0), < 75% of reads aligned to genome, or CG dinucleotide odds ratio > 0.5.

Retro-seq sample preparation—Mice (8–12-week-old) were anaesthetized (with 2%-5% isoflurane for induction and 0.8%–2% for maintenance), and mounted on a stereotaxic frame (David Kopf Instruments) with heating pad placed underneath. We bilaterally injected a volume of 200 nL containing retrograde tracer viruses into dPAG, lPAG, or MPOA of using a pulled glass capillary (World Precision Instruments) by pressure injection at a flow rate of 100 nL/min (Micoro4 controller, World Precision Instruments; Nanojector II, Drummond Scientific). The combination of retrogradely transported virus and mice used were as follows: 1) rAAV2-retro-EF1a-Cre (Tervo et al., 2016) in a Cre-reporter mouse line (*Ai14*$^{Tg/+}$ or *Ai110*$^{Tg/+}$) and 2) HSV1-LS1L-EYFP (rHSV; MGH Vector Core) in *Vglut2*$^{Cre/+}$ or *Esr1*$^{Cre/+}$ (Figure S1E). Stereotactic injection coordinates of dPAG, lPAG, and MPOA were obtained from the Paxinos and Franklin atlas (Franklin and Paxinos, 2008) (AP: −4.72, ML: ± 0.12, DV: −2.0 mm for dPAG; AP: −4.48, ML: ± 0.45, DV: −2.65 mm for lPAG; AP: 0.022, ML: ± 0.35, DV: −5.37 mm for MPOA). After a 2–3 week viral incubation, retrogradely labeled single cells were manually dissected from VMHvl, isolated by FACS, and subjected to scRNA-seq using SMART-seq as described above.

Single-cell isolation, cDNA amplification and library construction for 10x scRNA-seq and Act-seq—We isolated single cells from the mouse brain as previously described (Wu et al., 2017) with some modifications. Briefly, 40-55 min after the resident-intruder assay (see 'Resident intruder assay' section below), adult (8–12-week-old) resident mice (2–4 brains collected at a time) were anaesthetized with isoflurane and transcardially perfused with cold NMDG-ACSF (adjusted to pH 7.3–7.4) containing $CaCl_2$ (0.5 mM), glucose (25 mM), HCl (92 mM), HEPES (20 mM), KCl (2.5 mM), kynurenic acid (1 mM), $MgSO_4$ (10 mM), $NaHCO_3$ (30 mM), NaH_2PO_4 (1.2 mM), NMDG (92 mM), sodium L-ascorbate (5 mM), sodium pyruvate (3 mM), thiourea (2 mM), bubbled with carbogen gas. The brain was sectioned at 350 μm using vibratome (VT1000S, Leica Microsystems) on ice, and the regions of interest were microdissected under a fluorescence dissecting microscope from two consecutive sections (−1.22 to −1.94 from Bregma (Franklin and Paxinos, 2008)). The microdissected tissues were accumulated in a microcentrifuge tube containing NMDG-ACSF with 30 μM actinomycin D on ice to suppress further immediate early gene (IEG) activation (Wu et al., 2017). Thereafter the sections were transferred to a new microcentrifuge tube for papain digestion (60 U/mL, Sigma, P3125-250MG; pre-activated at 34°C for 30 min) in Trehalose-HEPES-ACSF (adjusted to pH 7.3–7.4) containing actinomycin D (15μM), $CaCl_2$ (2 mM), L-cysteine (2.5 mM), EDTA (0.5mM), glucose (25 mM), HEPES (20 mM), KCl (2.5 mM), kynurenic acid (1 mM), $MgSO_4$ (2 mM), NaCl (92 mM), $NaHCO_3$ (30 mM), NaH_2PO_4 (1.2 mM), trehalose (2.5% w/v) and gently carbogenated. During incubation for enzymatic digestion (~55 min at room temperature), the solution was mixed by gently pipetting a few times every 5-10 min. After incubation, the solution was exchanged with cold Trehalose-HEPES-ACSF containing Egg White/BSA ovomucoid inhibitor (3 mg/mL, Worthington, OI-BSA) and DNase I (25 U/mL, Thermo scientific, 90083). The tissue pieces were dissociated into single cells by gentle, successive trituration through Pasteur pipettes with polished tip openings of 600-μm, 300-μm and 150-

μm diameters. After trituration and filtering through a 40 μm cell strainer, single cells were pelleted at 300 g for 5 min at 4°C, the supernatant was carefully removed, the cells resuspended with cold Trehalose-HEPES-ACSF, and filtered through a 20 μm filter. Cells were pelleted again at 300 g for 5 min at 4°C and resuspended with Resuspension-HEPES-ACSF containing BSA (0.05%), $CaCl_2$ (2 mM), glucose (25 mM), HEPES (20 mM), KCl (2.5 mM), kynurenic acid (1 mM), $MgSO_4$ (1 mM), NaCl (117 mM), $NaHCO_3$ (30 mM), NaH_2PO_4 (1.2 mM) (osmolarity verified to be within 10 mOsm of Trehalose-HEPES-ACSF). After manually determining the cell concentration using a hemocytometer, suspensions were further diluted to desired concentrations (300–1,000 cells/μL) if necessary. The appropriate volume of reverse transcription (RT) mix was added in order to target 6,000–10,000 cells recovered, and loaded into the chip. The Chromium Single Cell 3′ Library & Gel Bead Kit v2 (PN- 120237), Chromium Single Cell 3′ Chip kit v2 (PN-120236), and Chromium i7 Multiplex Kit (PN-120262) were used for all downstream RT, cDNA amplification (11 PCR cycles), and library preparation as instructed by the manufacturer (Chromium Single Cell 3′ Reagents Kits v2 User Guide).

10x Sequencing data processing and quality control—Libraries were sequenced on an Illumina HiSeq4000 or NovaSeq6000 (paired-end with read lengths of 150 nt) and Illumina sequencing reads were aligned to the mouse pre-mRNA reference transcriptome (mm10) using the 10x Genomics CellRanger pipeline (version 3.0.2) with the default parameters. The median read depth per cell was 66,319 reads, yielding 2,998/1,309 genes (median), and 6,096/2,238 unique molecular identifier (UMI) counts (median) per neuronal/non-neuronal cell, respectively (Figure S4C, J, K). Cells that met any one of the following criteria were filtered out for downstream processing in each 10x run: < 600 detected genes (for UMI count >0), >30,000 UMI counts (potential multiplets), or the proportion of the UMI count attributable to mitochondrial genes was greater than 15%. Doublets were further removed by first classifying cells into broad cell classes (neuronal versus non-neuronal) based on the co-expression of any pair of their marker genes (*Stmn2* for neurons; *Cldn5* for endothelial cells; *C1qc* for microglia; *Opalin* for oligodendrocytes; *Gja1* for astrocytes; *Pdgfra* for OPCs; *Mustn1* for mural cells; see Figure S4A).

seqFISH data generation—seqFISH data in VMH and its surrounding area were generated using the seqFISH+ protocol with some modifications (Eng et al., 2019). In brief, primary probes of 30-nt sequences of each gene were extracted using the exons from within the CDS region and BLASTed against the mouse transcriptome to ensure specificity. A minimum of 15 to a maximum of 48 primary probes were designed for each targeted gene. The number of probes per gene was dependent on the length of the CDS region. The same 15-nt readout probes were used as in previous study (Shah et al., 2018). The reverse complements of these readout probe sequences were included as four binding sites in the primary probes to increase signals. The final primary probes were ordered as a complex oligopool from Twist Bioscience and were constructed as previously described (Eng et al., 2019; Shah et al., 2018).

Adult (8–12-week-old) C57BL/6N male and female mice were perfused with perfusion buffer (10U/mL heparin, 0.5% $NaNO_2$ (w/v) in 1X PBS at 4°C). Mice were then perfused

with fresh 4% paraformaldehyde (PFA; Thermo Scientific, 28908)/1XPBS buffer (Invitrogen, AM9624). The brain was dissected and immediately placed in a 4% PFA buffer for 2 h at room temperature. The brain was then immersed in 30% RNase-free Sucrose (Amresco, 0335) in 1X PBS for 48 h at 4°C until the brain sank to the bottom of the tube. The brain was then embedded in OCT and cryo-sectioned (15 μm thick), and 4–5 sections including VMHvl were collected per mouse (consecutive sections were ~100 μm apart along A-P axis; 20 sections in total). The sections were stored at −80°C until use.

For seqFISH data acquisition, sections were prepared as described with modifications (Eng et al., 2019). In brief, sections were permeabilized in 70% ethanol for 18 h at 4°C, and further permeabilized by the addition of RNase-free 8% SDS (Ambion, AM9822) in 1X PBS for 15 min at room temperature. Tissue slices were then rinsed multiple times with 70% ethanol to remove SDS. Primary probe hybridization buffer composed of 40% Formamide (Invitrogen, AM9344), 2X SSC (Invitrogen, 15557-044) 10% (w/v) Dextran Sulfate (Sigma, D8906), and ~5nM/primary probes were hybridized to the tissue slices by spreading out the hybridization buffer solution with a coverslip. The hybridization was allowed to proceed for ~36 h at 37°C incubator in a humid chamber. After primary probe hybridization, the tissue slices were washed with 40% Wash Buffer (40% WB: 2x SSC, 40% Formamide (v/v), 0.1% Triton X-100 (Sigma, D8906) at 37°C for 30 min. After rinsing with 2X SSC 3 times and 1X PBS once, the sample was subjected to 0.1mg/mL Acryoloyl-X SE (Thermo Fisher, A20770) in 1X PBS treatment for 30 min at room temperature. Next, the tissue slices were incubated with 4% acrylamide (1:19 crosslinking) hydrogel solution in 2X SSC for 30 min at room temperature. Then the hydrogel solution was aspirated and 30 μL of 4% hydrogel solution containing 0.05% APS and 0.05% TEMED in 2x SSC was dropped onto the tissue slice and sandwiched by Gel-Slick functionalized coverslip. The samples were transferred to 4°C in a homemade nitrogen gas chamber for 30 min before transferring to 37°C for 2.5 h to complete polymerization. After polymerization, the hydrogel embedded tissue slices were cleared with digestion buffer as previously described (Eng et al., 2019). After digestion, the tissue slices were rinsed with 2X SSC multiple times and subjected to 0.1mg/mL Label-X modification for 45 min at 37°C. The tissue slices were reembedded in hydrogel solution as in the previous step, and were gelated at 37°C incubator for 2 h for stronger stabilization before subsequent multiple rounds of imaging. Imaging platform and automated fluidics delivery system were similar to those previously described (Eng et al., 2019).

Each round of imaging contained the 405 channel, which included the DAPI stain of the cell. The DAPI images from all the rounds of hybridization were aligned to the first image using a 3D phase correlation algorithm. Tissue background and auto-fluorescence were then removed by dividing the initial background with the fluorescence images. To correct for the non-uniform background, a flat field correction was applied by dividing the normalized background illumination with each of the fluorescence images while preserving of the intensity profile of the fluorescent points. The background signal was then subtracted using the ImageJ rolling ball background subtraction algorithm with a radius of 3 pixels. For semi-automatic cell segmentation, images of sections stained with Nissl and DAPI were trained with Ilastik (Sommer et al., 2011), an interactive supervised machine learning toolkit, to output probability maps, which were then used in the Multicut (Beier et al., 2017) tool to produce volumetric labeled cells. The labeled cells were filtered by size and region in

MATLAB. For image analysis, potential mRNA transcript signals were located by finding the local maxima in the processed image above a predetermined pixel threshold. The transcript spots were assigned to the corresponding labeled cells according to location. For more details regarding the seqFISH method, see (Shah et al., 2016).

Resident intruder assay—The resident intruder assay was performed as previously described (Hong et al., 2014; Lee et al., 2014). Briefly, all experimental male mice ("residents") were had been individually housed for two weeks and received sexually experience (for 1–3 days), except residents for "Social Fear" assays using group-housed (sexually naive) males. Before testing, the resident mice were pre-screened for baseline aggression and sexual behavior using resident-intruder testing sessions (~3 times). On the experimental day, residents were transported in their home cage to a novel behavioral testing room (under infrared light) where they acclimated for 5–15 min. An unfamiliar group-housed BALB/c mouse ("intruders"; types of intruders listed in Figure 5A) was then placed in the home cage of the resident, and residents were allowed to freely interact with it for ~10 min. For"M-M CI" (male-male close investigation only) and "M-F CI" (male-female close investigation only) tests, intruders were either dangled by the tail or placed inside a wire-mesh barrier (an inverted pencil cup) when introduced into the resident's home cage, so that resident behavior was restricted to investigation and not allowed to proceed to the consummatory phases (fighting or copulation). "Control" indicates home-caged animals moved to the behavior testing room, but not otherwise manipulated. "Plain" indicated animals taken directly from their home cage in the housing room. In the cases of "Aggression" and "Mating," only resident mice that were actively engaging in either aggression or sexual behaviors throughout the session were used for subsequent 10x scRNA-seq experiments. For "Social Fear," experimental mice were introduced to the home cage of singly housed male aggressors for ~10 min to allow them to be attacked (more than 10 times), and then returned into their home cage. For female mice, sexually naive and group-housed mice were used as residents for the "Plain" category (n = 5,935 cells from 4 mice), "Control" (n = 7,964 cells from 4 mice), or social interaction tests with male intruders (n = 12,267 cells from 8 mice; most of them were unreceptive), respectively.

Retrograde labeling with *c-fos* immunohistochemistry—For retrograde labeling, 200 nL of HSV-LS1L-mCherry or HSV-LS1L-EYFP (MGH Vector Core) was injected into either dPAG or MPOA of adult $Esr1^{cre/+}$ mice (8–12 weeks old). ~3 weeks after injections, mice were tested in resident-intruder assays (3 main categories: male-male aggression, male-female mounting, and male-female close investigation only). 80–90 min after the resident-intruder assay, mice were euthanized via transcardiac perfusion with 4% PFA, and serial sections generated using a cryostat. To visualize *c-fos,* sections were stained with goat anti-*c-fos* antibody (Santa Cruz Biotechnology, sc-52-g, 1:500) overnight at 4°C, followed by Alexa-647 donkey anti-goat (Invitrogen, A-21447,1:1000) at room temperature for 3 h. For manual quantification of fluorescent cells, images were analyzed using Fiji/ImageJ as follows: 1) a manually set threshold was applied (to the point that background autofluorescence was not visible), 2) watershed-based segmentation and particle analysis ("analyze particles" function) was used for the identification of individual cells (with size

and circularity limited to an appropriate range), and 3) automatic particle counting was visually inspected for accuracy and manually adjusted.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data were processed and analyzed using MATLAB, R, Prism 6, ImageJ, and Excel. The following statistical analyses were conducted: ANOVAs followed by Bonferroni post hoc tests to compare 1) the expression of *Fos* induced by each behavioral paradigm with control across clusters (Figures 5B and 7A), and 2) the proportion of double$^+$ (*Fos* & *rHSV*) cells between aggression and control injected in dPAG and MPOA (Figure 7D); Fisher's exact tests to see 1) which retrogradely labeled SMART-seq classes/clusters from a given projection were significantly different from the proportions of retrogradely labeled cells from other projection (Figures 3B and 3C), 2) what kinds of joint CCA clusters were significantly different from the expected proportions of cells in a given sex (Figure 6B), and 3) how differently cells from sex-specific clusters were located between anterior versus posterior VMHvl (Figure 6E); a likelihood ratio test (McDavid et al., 2013) was used to calculate adjusted *p*-values (FindMarkers function in Seurat; after Bonferroni correction using all genes in the dataset) for determining the significance values of differential IEG expression compared to controls (Figure S5B). Unpaired t tests were used to compute *p*-values for 1) the differential expressions of genes in seqFISH (Figure 6F), and 2) different proportion of double$^+$ (*Fos* & *rHSV*) cells between M-F social interaction tests (mating + M-F CI tests) and control both in dPAG and MPOA injections (Figure 7D). The n value and statistically significant effects are reported in each figure/figure legend. The significance threshold was held at $\alpha = 0.05$ ($p > 0.05$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$).

SMART-seq data analysis and clustering—Data analysis for the SMART-seq dataset (including clustering, dendrogram construction and differential gene expression) was performed using a pre-release version of R package scrattch.hicat (https://github.com/AllenInstitute/scrattch.hicat), as described previously (Tasic et al., 2018). In brief, all QC qualified cells were clustered using high variance gene selection, dimensionality reduction (using default PCA mode), dimension filtering, and Jaccard-Louvain or hierarchical (Ward) clustering. This process was repeated within each resulting cluster until no more child clusters met differential gene expression or cluster size termination criteria. The entire clustering procedure was repeated 100 times using 80% of all cells sampled at random, and the frequency with which cells co-cluster was used to generate a final set of clusters (Figure S2A), again subject to differential gene expression and cluster size termination criteria. Pairwise differential gene expression was performed using the de_score function in the scrattch.hicat package based on $\log_2$(CPM+1) values. The cell type tree was built using the build_dend function in scrattch.hicat based on cluster median expression profiles of marker genes. Neighboring clusters in the cell type taxonomy with shared key markers are defined as types within the subclass, and clusters are named for key markers with a suffix for further discrimination.

To evaluate the clustering robustness, we compared the original SMART-seq cluster membership (Figure 1C; based on scratch.hicat iterative clustering) with different data

processing and clustering algorithms (either Ward's hierarchical clustering (Figure S2D) or graph-based clustering by Seurat (Figure S2E)). For Ward's hierarchical clustering (Ward, 1963), gene counts across 45,768 exons in single cells were log-normalized. Because dimensionality reduction on this dataset was spatially structured by cell sequencing depth, we normalized for sequencing depth per cell by regressing out the log of the sum of the exon counts per cell. After normalization, PCA followed by t-SNE no longer arranged cells by sequencing depth. scipy.cluster.hierarchy.linkage was then used to perform hierarchical clustering using the Ward variance minimization metric on the normalized counts, and the hierarchical clustering was cut at 38 clusters. For graph-based clustering, SMART-seq gene counts were loaded into Seurat. After log-normalization and data scaling (using default parameters), the top 2,000 highly variable genes were used (FindVariableGenes function; selection.method = 'vst') to identify 44 clusters (FindClusters function; using 40 PCs given the highly variable genes and a resolution of 4.5).

All retrogradely labeled SMART-seq cells were subjected to the same experimental and data processing and clustering with all other quality control-passed single-cells. Clustering was performed blinded to the experimental source of retrogradely labeled cells.

10x Data analysis and clustering—All downstream analysis of Act-seq 10x scRNA-seq data was performed with R package Seurat (v3.0.3) (Butler et al., 2018; Satija et al., 2015; Stuart et al., 2019), using default parameters unless specified (https://satijalab.org/seurat/). Briefly, after initial cell filtering, genes expressed in fewer than 0.1% of total cells in each experiment were also removed, and all 10x gene expression matrices from each experiment were loaded and merged into Seurat. Gene expressions of each cell were normalized by total number of molecules, multiplied by a scale factor (10,000) per cell, and log-transformed (NormalizeData function). Then the expressions of each gene were scaled (its mean/variance across cells is 0 and 1, respectively) and the proportion of mitochondrial UMIs was regressed out (ScaleData function). Next, highly variable genes were identified (FindVariableGenes function; top 2,000 genes with the highest standardized variance selected by selection.method = 'vst') and used as input for dimensionality reduction via principal component analysis (PCA) after removing sex-specific genes (*Ddx3y*, *Eif2s3y*, *Uty*, *Kdm5d*, *Xist*, *Tsix*), immediate early genes (e.g., *Fos*, *Fosl2*, *Junb*, *Egr1*, *Arc*, *Homer1*; 139 genes in total from (Wu et al., 2017)), 30 retro-virus-induced genes (e.g., *B2m*, *Bst2*, *Oasl2*, *Ifit1*), and 1,000 noise-sensitive genes (high abundance genes sensitive to technical noise; see also Table S2). The resulting PCs were then used as input for clustering analysis (FindClusters function). Specifically, we performed iterative rounds of clustering and cell selection (Figures S4A–S4H). For the initial 149,663 cells, we used 40 PCs to classify 35 unique clusters (a resolution of 0.6) to separate neuronal cells from non-neuronal cells. We then selected the putative neuronal clusters (#0–21) for the 2nd round of clustering analysis. 56 neuronal clusters were identified using 40 PCs (a resolution of 2.0) from 78,476 neurons to separate glutamatergic VMH-in (inside VMH) neurons from GABAergic VMH-out (outside VMH) neurons. The final round of clustering analysis was performed given 41,385 VMH neurons (#1–23), and 29 VMH clusters were identified using 40 PCs (a resolution of 1.6; Figures 4C–4E). Clustering robustness for VMH clusters was evaluated by varying the parameters (number of PCs: 30–60; resolution: 1–2), and cluster identification was robust

across the range of PCs and resolutions. For VMH clusters, top 10 differentially expressed genes were computed by FindAllMarkers function (min.pct = 0.25, logfc.threshold = 0.25; Figure S4L).

To compare the expression level of *Fos* in each resident-intruder assay with control in whole VMHvl clusters (Figure 5B; 14,763 cells from 10x VMH cluster #2–17; only cells from male mice were included), we calculated *p*-values after multiple comparison correction across behaviors and clusters (using 3-way ANOVA and Bonferroni post hoc test). We also performed the same analysis using only cells from M-M social interaction tests ("M-M"; aggression and two M-M CI tests) or M-F social interaction tests ("M-F"; mating and two M-F CI tests) samples in the 6 Esr1 clusters (#3-8; Figure 5C; 3,972,11,094, and 4,865 cells from control, M-M, and M-F, respectively). For 7 selected VMHvl clusters (Esr1_4, Esr1_5, Esr1_6, Esr1_7, Nup62cl, Dlk1_1, and Dlk1_4; showing strong *Fos* inductions in Figure 5B), 1) pairwise comparisons between each social behavior exhibiting significant differences in *Fos* expression (Figure S5A; after multiple comparison correction), and 2) differential IEG expression analysis (for 139 IEGs; see Table S2) using Seurat (a likelihood ratio test (McDavid et al., 2013) by FindMarkers function) to identify significantly upregulated (average fold change > 2 and FDR < 0.05) IEGs in each behavior compared to control (Figure S5B) were performed. For each VMHvl cluster, the number of IEGs significantly induced (average fold change > 2 and FDR < 0.05) during each social behavior compared to control was also computed (Figure 5E).

seqFISH Data analysis and clustering—The smFISH counts for each gene were normalized by Z-scoring within each brain section, and for imaging ROIs that covered the same portion of VMHvl, their spatial coordinates were integrated into a common coordinate system. Based on Nissl-stained images, VMH and its subdivisions (VMHvl or VMHc) were identified, and cells were assigned to one of the major anatomic regions (VMHvl, VMHc and VMH-out). For 4,686 cells in VMHvl, their gene expression matrix was loaded into Seurat, PCA and clustering was performed (using 40 PCs and a resolution of 3.4 in FindClusters). Among initial 29 clusters, we filtered out 2 clusters that showed very low expression levels of all genes and are mostly located nearby the boundary of the ROIs. The dendrogram for final 27 clusters (4,497 cells) was constructed based on the average distance between clusters in PCA space (BuildClusterTree function; Figure 2B).

To establish correspondence between scRNA-seq and seqFISH clusters, we computed the pairwise Pearson's correlation coefficient between the expression patterns of 58 seqFISH genes in SMART-seq (3,824 cells) and seqFISH (4,497 cells) VMHvl dataset (using corrcoef function in MATLAB). We then identified the seqFISH clusters that were significantly correlated with SMART-seq clusters ($p < 0.05$), and whose correlation coefficients were greater than the median value of bootstrap replicates ($r > 0.35$). Among these, the seqFISH clusters most strongly correlated with each scRNA-seq cluster are marked by red squares (Figure 2C). 23 out of 27 seqFISH clusters corresponded to one or multiple SMART-seq clusters, and the pairwise correlation between these seqFISH clusters and scRNA-seq subclasses showed a clear diagonal relationship (Figure 2C).

MetaNeighbor analysis—To identify the set of genes that highly predict each of the 7 major Act-seq clusters identified in VMHvl (Esr1_4, Esr1_5, Esr1_6, Esr1_7, Nup62cl, Dlk1_1, and Dlk1_4), we calculated AUROC (Area Under Receiver-Operator Characteristics Curve) scores using MetaNeighbor as described previously (Crow et al., 2018; Paul et al., 2017). Briefly, we inputted ~500 HGNC gene families (Paul et al., 2017) and a gene expression matrix for each of the 7 clusters (7,632 cells only from male mice), with their cell type identities (and a single experiment labeling only) to MetaNeighbor function. Output AUROC scores reflect the probability of a correct assignment of cell types based on a given set of genes, when making a single binary choice, using randomized labels (AUROC ~0.5) as control (see Table S3).

Canonical correlation analysis (CCA) with SMART-seq and 10x datasets—To integrate the two different scRNA-seq datasets, their shared sources of variation were identified by Seurat CCA alignment (Butler et al., 2018; Stuart et al., 2019). Because the 10x data contained a large number of contaminating cells outside of VMH, in order to make a fair comparison, only cells from VMH clusters were considered from each dataset (3,824 cells from SMART-seq clusters #20–46; 41,385 cells from 10x). Specifically, after data preprocessing (log-normalization, data scaling for each gene, and detecting variable genes) on SMART-seq data in Seurat, we selected the top 2,000 genes with the highest dispersion (variance to mean ratio) from each dataset. We took the union of these two resulting gene sets (2,338 in total) as input to identify the pairwise correspondences between single cells across two datasets, called anchors (FindIntegrationAnchors function). We then used these anchors to integrate the two datasets together (IntegrateData function), performed a joint clustering on these aligned embeddings (FindClusters function; resolution of 1.9), yielding 31 clusters (referred to as “Joint Clusters (CCA)”; Figures 6A and S7C). Lastly, differentially expressed genes that were conserved across datasets were identified (FindConservedMarkers function), and a dendrogram was built using build_dend function in scrattch.hicat based on cluster median expression profiles of marker genes (Figure 6I).

DATA AND CODE AVAILABILITY

Data analysis and visualization software—The accession number for the scRNA-seq data reported in this paper is https://doi.org/10.17632/ypx3sw2f7c.1 (Mendeley Data). Analysis and visualization of transcriptomic data were performed using MATLAB (version 2017a) and R v.3.3.0 and greater, assisted by the Rstudio IDE (Integrated Development Environment for R v.1.1.442; https://rstudio.com/) as well as the following R packages: cowplot v.0.9.2 (https://rdrr.io/cran/cowplot/), dendextend v.1.5.269, dplyr v.0.7.4 (https://dplyr.tidyverse.org/), feather v0.3.1 (https://rdrr.io/cran/feather/), ggbeeswarm v.0.6.0 (https://cran.r-project.org/web/packages/ggbeeswarm/index.html), ggExtra v.0.8 (https://rdrr.io/cran/ggExtra/), ggplot2 v.2.2.170, ggrepel v.0.7.0 (https://cran.r-project.org/web/packages/ggrepel/vignettes/ggrepel.html), googlesheets v.0.2.2 (https://cran.r-project.org/web/packages/googlesheets/vignettes/basic-usage.html), gridExtra v.2.3 (https://cran.r-project.org/web/packages/gridExtra/index.html), Hmisc v.4.1-1 (https://cran.r-project.org/web/packages/Hmisc/index.html), igraph v.1.2.1 (https://www.rdocumentation.org/packages/igraph/versions/1.2.1), limma v.3.30.1366,71, Matrix v.1.2-12 (https://rdrr.io/rforge/Matrix/), matrixStats v.0.53.1 (https://cran.rstudio.com/web/

packages/matrixStats/index.html), pals v.1.5 (https://rdrr.io/cran/pals/), purrr v.0.2.4 (https://purrr.tidyverse.org/), pvclust v.2.0-0 (http://stat.sys.i.kyoto-u.ac.jp/prog/pvclust/), reshape2 v.1.4.2 (https://www.statmethods.net/management/reshape.html), Rphenograph v.0.99.1 (https://rdrr.io/github/JinmiaoChenLab/Rphenograph/), Rtsne v.0.14. (https://cran.r-project.org/web/packages/Rtsne/citation.html), Seurat v.3.0.3, viridis v.0.5.0 (https://rdrr.io/cran/viridisLite/man/viridis.html), and xlsx v.0.5.7 (https://cran.r-project.org/web/packages/xlsx/index.html).

Code availability—An R package for iterative clustering (scrattch.hicat) is available on GitHub at https://github.com/AllenInstitute/scrattch.hicat. All MATLAB and R scripts used in this manuscript are available on reasonable request.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank A. Jones and C. Koch for support at the Allen Institute for Brain Sciences during the writing of this manuscript, J.-S. Chang for cell counting, Y. Huang for genotyping, G. Mancuso for administrative assistance, C. Chiu for lab management, S. Diamond for assistance with FACS, the Single Cell Profiling and Engineering Center (SPEC) in the Beckman Institute at Caltech for initial help for 10x scRNA-seq experiments, S. Pease for assistance with transgenic mouse strains, J. Costanza for mouse colony management, members of the Anderson laboratory for helpful comments on this project, and an anonymous reviewer for suggesting social fear testing in group-housed mice. This work was supported by US National Institutes of Health (NIH) BRAIN Initiative grants U01MH105982 and U19MH114830 to H.Z. and D.J.A. and NIH grants MH070053 and TR01 OD024686 to D.J.A. and L.C., respectively. D.-W.K. was supported by a Howard Hughes Medical Institute International Student Research Fellowship. D.J.A. is an investigator of the Howard Hughes Medical Institute.

REFERENCES

Beier T, Pape C, Rahaman N, Prange T, Berg S, Bock DD, Cardona A, Knott GW, Plaza SM, Scheffer LK, et al. (2017). Multicut brings automated neurite segmentation closer to human performance. Nat. Methods 14, 101–102. [PubMed: 28139671]

Bravo R (1990). Growth factor-responsive genes in fibroblasts. Cell Growth Differ. 1, 305–309. [PubMed: 2126191]

Butler A, Hoffman P, Smibert P, Papalexi E, and Satija R (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol 36, 411–420. [PubMed: 29608179]

Cachero S, Ostrovsky AD, Yu JY, Dickson BJ, and Jefferis GSXE (2010). Sexual dimorphism in the fly brain. Curr. Biol 20, 1589–1601. [PubMed: 20832311]

Campbell JN, Macosko EZ, Fenselau H, Pers TH, Lyubetskaya A, Tenen D, Goldman M, Verstegen AMJ, Resch JM, McCarroll SA, et al. (2017). A molecular census of arcuate hypothalamus and median eminence cell types. Nat. Neurosci 20, 484–496. [PubMed: 28166221]

Canteras NS, Simerly RB, and Swanson LW (1994). Organization of projections from the ventromedial nucleus of the hypothalamus: a Phaseolus vulgaris-leucoagglutinin study in the rat. J. Comp. Neurol 348, 41–79. [PubMed: 7814684]

Chen P, and Hong W (2018). Neural Circuit Mechanisms of Social Behavior. Neuron 98, 16–30. [PubMed: 29621486]

Chen R, Wu X, Jiang L, and Zhang Y (2017). Single-Cell RNA-Seq Reveals Hypothalamic Cell Diversity. Cell Rep. 18, 3227–3241. [PubMed: 28355573]

Cheung CC, Kurrasch DM, Liang JK, and Ingraham HA (2013). Genetic labeling of steroidogenic factor-1 (SF-1) neurons in mice reveals ventromedial nucleus of the hypothalamus (VMH) circuitry

beginning at neurogenesis and development of a separate non-SF-1 neuronal cluster in the ventrolateral VMH. J. Comp. Neurol 521, 1268–1288. [PubMed: 22987798]

Choi GB, Dong HW, Murphy AJ, Valenzuela DM, Yancopoulos GD, Swanson LW, and Anderson DJ (2005). Lhx6 delineates a pathway mediating innate reproductive behaviors from the amygdala to the hypothalamus. Neuron 46, 647–660. [PubMed: 15944132]

Correa SM, Newstrom DW, Warne JP, Flandin P, Cheung CC, Lin-Moore AT, Pierce AA, Xu AW, Rubenstein JL, and Ingraham HA (2015). An estrogen-responsive module in the ventromedial hypothalamus selectively drives sex-specific activity in females. Cell Rep. 10, 62–74. [PubMed: 25543145]

Crow M, Paul A, Ballouz S, Huang ZJ, and Gillis J (2018). Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. Nat. Commun 9, 884. [PubMed: 29491377]

Daigle TL, Madisen L, Hage TA, Valley MT, Knoblich U, Larsen RS, Takeno MM, Huang L, Gu H, Larsen R, et al. (2018).ASuite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality. Cell 174, 465–480 e422. [PubMed: 30007418]

Dhillon H, Zigman JM, Ye C, Lee CE, McGovern RA, Tang V, Kenny CD, Christiansen LM, White RD, Edelstein EA, et al. (2006). Leptin directly activates SF1 neurons in the VMH, and this action by leptin is required for normal body-weight homeostasis. Neuron 49, 191–203. [PubMed: 16423694]

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. [PubMed: 23104886]

Economo MN, Viswanathan S, Tasic B, Bas E, Winnubst J, Menon V, Graybuck LT, Nguyen TN, Smith KA, Yao Z, et al. (2018). Distinct descending motor cortex pathways and their roles in movement. Nature 563, 79–84. [PubMed: 30382200]

Eng CL, Lawson M, Zhu Q, Dries R, Koulena N, Takei Y, Yun J, Cronin C, Karp C, Yuan GC, and Cai L (2019). Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. Nature 568, 235–239. [PubMed: 30911168]

Falkner AL, Dollar P, Perona P, Anderson DJ, and Lin D (2014). Decoding ventromedial hypothalamic neural activity during male mouse aggression. J. Neurosci 34, 5971–5984. [PubMed: 24760856]

Franklin KBJ, and Paxinos G (2008). The mouse brain in stereotaxic coordinates (Academic Press).

Greenberg ME, and Ziff EB (1984). Stimulation of 3T3 cells induces transcription of the c-fos proto-oncogene. Nature 311, 433–438. [PubMed: 6090941]

Hahn JD, Sporns O, Watts AG, and Swanson LW (2019). Macroscale intrinsic network architecture of the hypothalamus. Proc. Natl. Acad. Sci. USA 116, 8018–8027. [PubMed: 30923123]

Hashikawa K, Hashikawa Y, Tremblay R, Zhang J, Feng JE, Sabol A, Piper WT, Lee H, Rudy B, and Lin D (2017a). Esr1^{+} cells in the ventromedial hypothalamus control female aggression. Nat. Neurosci 20, 1580–1590. [PubMed: 28920934]

Hashikawa Y, Hashikawa K, Falkner AL, and Lin D (2017b). Ventromedial Hypothalamus and the Generation of Aggression. Front. Syst. Neurosci 11, 94. [PubMed: 29375329]

Hong W, Kim DW, and Anderson DJ (2014). Antagonistic control of social versus repetitive self-grooming behaviors by separable amygdala neuronal subsets. Cell 158, 1348–1361. [PubMed: 25215491]

Hrvatin S, Hochbaum DR, Nagy MA, Cicconet M, Robertson K, Cheadle L, Zilionis R, Ratner A, Borges-Monroy R, Klein AM, et al. (2018). Single-cell analysis of experience-dependent transcriptomic states in the mouse visual cortex. Nat. Neurosci 21, 120–129. [PubMed: 29230054]

Ishii KK, Osakada T, Mori H, Miyasaka N, Yoshihara Y, Miyamichi K, and Touhara K (2017). A Labeled-Line Neural Circuit for Pheromone-Mediated Sexual Behaviors in Mice. Neuron 95, 123–137 e128. [PubMed: 28648498]

Jorgenson LA, Newsome WT, Anderson DJ, Bargmann CI, Brown EN, Deisseroth K, Donoghue JP, Hudson KL, Ling GS, MacLeish PR, et al. (2015). The BRAIN Initiative: developing technology to catalyse neuroscience discovery. Philos. Trans. R. Soc. Lond. B Biol. Sci 370, 20140164. [PubMed: 25823863]

Kennedy A, Asahina K, Hoopfer E, Inagaki H, Jung Y, Lee H, Remedios R, and Anderson DJ (2014). Internal States and Behavioral Decision-Making: Toward an Integration of Emotion and Cognition. Cold Spring Harb. Symp. Quant. Biol 79, 199–210. [PubMed: 25948637]

Kohl J, Babayan BM, Rubinstein ND, Autry AE, Marin-Rodriguez B, Kapoor V, Miyamishi K, Zweifel LS, Luo L, Uchida N, and Dulac C (2018). Functional circuit architecture underlying parental behaviour. Nature 556, 326–331. [PubMed: 29643503]

Krause WC, and Ingraham HA (2017). Origins and Functions of the Ventrolateral VMH: A Complex Neuronal Cluster Orchestrating Sex Differences in Metabolism and Behavior. Adv. Exp. Med. Biol 1043, 199–213. [PubMed: 29224096]

Kunwar PS, Zelikowsky M, Remedios R, Cai H, Yilmaz M, Meister M, and Anderson DJ (2015). Ventromedial hypothalamic neurons control a defensive emotion state. eLife 4, e06633.

Kurrasch DM, Cheung CC, Lee FY, Tran PV, Hata K, and Ingraham HA (2007). The neonatal ventromedial hypothalamus transcriptome reveals novel markerswith spatially distinct patterning. J. Neurosci 27,13624–13634. [PubMed: 18077674]

Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, and Carey VJ (2013). Software for computing and annotating genomic ranges. PLoS Comput. Biol 9, e1003118. [PubMed: 23950696]

Lee H, Kim DW, Remedios R, Anthony TE, Chang A, Madisen L, Zeng H, and Anderson DJ (2014). Scalable control of mounting and attack by Esr1+ neurons in the ventromedial hypothalamus. Nature 509, 627–632. [PubMed: 24739975]

Li H, Horns F, Wu B, Xie Q, Li J, Li T, Luginbuhl DJ, Quake SR, and Luo L (2017). Classifying Drosophila Olfactory Projection Neuron Subtypes by Single-Cell RNA Sequencing. Cell 171, 1206–1220 e1222. [PubMed: 29149607]

Lin D, Boyle MP, Dollar P, Lee H, Lein ES, Perona P, and Anderson DJ (2011). Functional identification of an aggression locus in the mouse hypothalamus. Nature 470, 221–226. [PubMed: 21307935]

Liu KS, and Sternberg PW (1995). Sensory regulation of male mating behavior in Caenorhabditis elegans. Neuron 14, 79–89. [PubMed: 7826644]

Lo L, Yao S, Kim DW, Cetin A, Harris J, Zeng H, Anderson DJ, and Weissbourd B (2019). Connectional architecture of a mouse hypothalamic circuit node controlling social behavior. Proc. Natl. Acad. Sci. USA 116, 7503–7512. [PubMed: 30898882]

Lovett-Barron M, Andalman AS, Allen WE, Vesuna S, Kauvar I, Burns VM, and Deisseroth K (2017). Ancestral Circuits for the Coordinated Modulation of Brain State. Cell 171, 1411–1423 e1417. [PubMed: 29103613]

Luiten PGM, ter Horst GJ, and Steffens AB (1987). The hypothalamus, intrinsic connections and outflow pathways tothe endocrine system in relation to the control of feeding and metabolism. Prog. Neurobiol 28, 1–54. [PubMed: 3547503]

Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M,Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 161, 1202–1214. [PubMed: 26000488]

Madisen L, Zwingman TA, Sunkin SM, Oh SW, Zariwala HA, Gu H, Ng LL, Palmiter RD, Hawrylycz MJ, Jones AR, et al. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci 13, 133–140. [PubMed: 20023653]

McDavid A, Finak G, Chattopadyay PK, Dominguez M, Lamoreaux L, Ma SS, Roederer M, and Gottardo R (2013). Dataexploration, quality control and testing in single-cell qPCR-based gene expression experiments. Bioinformatics 29, 461–467. [PubMed: 23267174]

McHenry JA, Otis JM, Rossi MA, Robinson JE, Kosyk O, Miller NW, McElligott ZA, Budygin EA, Rubinow DR, and Stuber GD (2017). Hormonal gain control of a medial preoptic area social reward circuit. Nat. Neurosci 20, 449–458. [PubMed: 28135243]

Mickelsen LE, Bolisetty M, Chimileski BR, Fujita A, Beltrami EJ, Costanzo JT, Naparstek JR, Robson P, and Jackson AC (2019). Single-cell transcriptomic analysis of the lateral hypothalamic area reveals molecularly distinct populations of inhibitory and excitatory neurons. Nat. Neurosci 22, 642–656. [PubMed: 30858605]

Millhouse OE (1973a). Certain ventromedial hypothalamic afferents. Brain Res. 55, 89–105. [PubMed: 4197323]

Millhouse OE (1973b). The organization of the ventromedial hypothalamic nucleus. Brain Res. 55, 71–87. [PubMed: 4713193]

Moffitt JR, Hao J, Wang G, Chen KH, Babcock HP, and Zhuang X (2016). High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. Proc. Natl. Acad. Sci. USA 113, 11046–11051. [PubMed: 27625426]

Moffitt JR, Bambah-Mukku D, Eichhorn SW, Vaughn E, Shekhar K, Perez JD, Rubinstein ND, Hao J, Regev A, Dulac C, et al. (2018). Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. Science 362, eaau5324. [PubMed: 30385464]

Moon YS, Smas CM, Lee K, Villena JA, Kim KH, Yun EJ, and Sul HS (2002). Mice lacking paternally expressed Pref-1/Dlk1 display growth retardation and accelerated adiposity. Mol. Cell. Biol 22, 5585–5592. [PubMed: 12101250]

Morgan JI, Cohen DR, Hempstead JL, and Curran T (1987). Mapping patterns of c-fos expression in the central nervous system after seizure. Science 237, 192–197. [PubMed: 3037702]

Morrell JI, and Pfaff DW (1982). Characterization of estrogen-concentrating hypothalamic neurons by their axonal projections. Science 217, 1273–1276. [PubMed: 7112131]

Musatov S, Chen W, Pfaff DW, Mobbs CV, Yang XJ, Clegg DJ, Kaplitt MG, and Ogawa S (2007). Silencing of estrogen receptor alpha in the ventromedial nucleus of hypothalamus leads to metabolic syndrome. Proc. Natl. Acad. Sci. USA 104, 2501–2506. [PubMed: 17284595]

Oh SW, Harris JA, Ng L, Winslow B, Cain N, Mihalas S, Wang Q, Lau C, Kuan L, Henry AM, et al. (2014). A mesoscale connectome of the mouse brain. Nature 508, 207–214. [PubMed: 24695228]

Paul A, Crow M, Raudales R, He M, Gillis J, and Huang ZJ (2017). Transcriptional Architecture of Synaptic Communication Delineates GABAergic Neuron Identity. Cell 171, 522–539 e520. [PubMed: 28942923]

Persson-Augner D, Lee YW, Tovar S, Dieguez C, and Meister B (2014). Delta-like 1 homologue (DLK1) protein in neurons of the arcuate nucleus that control weight homeostasis and effect of fasting on hypothalamic DLK1 mRNA. Neuroendocrinology 100, 209–220. [PubMed: 25342302]

Pfaff DW (2017). How the Vertebrate Brain Regulates Behavior: Direct from the Lab (Harvard University Press).

Pfaff DW, and Sakuma Y (1979a). Deficit in the lordosis reflex of female rats caused by lesions in the ventromedial nucleus of the hypothalamus. J. Physiol 288,203–210. [PubMed: 469716]

Pfaff DW, and Sakuma Y (1979b). Facilitation of the lordosis reflex of female rats from the ventromedial nucleus of the hypothalamus. J. Physiol 288, 189–202. [PubMed: 469715]

Picelli S, Björklund AK, Faridani OR, Sagasser S, Winberg G, and Sandberg R (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat. Methods 10, 1096–1098. [PubMed: 24056875]

Remedios R, Kennedy A, Zelikowsky M, Grewe BF, Schnitzer MJ, and Anderson DJ (2017). Social behaviour shapes hypothalamic neural ensemble representations of conspecific sex. Nature 550, 388–392. [PubMed: 29052632]

Roeling TA, Veening JG, Kruk MR, Peters JP, Vermelis ME, and Nieuwenhuys R (1994). Efferent connections of the hypothalamic “aggression area” in the rat. Neuroscience 59, 1001–1024. [PubMed: 8058117]

Romanov RA, Zeisel A, Bakker J, Girach F, Hellysaz A,Tomer R, Alpar A, Mulder J, Clotman F, Keimpema E, et al. (2017). Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. Nat. Neurosci 20, 176–188. [PubMed: 27991900]

Sakurai K, Zhao S, Takatoh J, Rodriguez E, Lu J, Leavitt AD, Fu M, Han B-X, and Wang F (2016). Capturing and Manipulating Activated Neuronal Ensembles with CANE Delineates a Hypothalamic Social-Fear Circuit. Neuron 92, 739–753. [PubMed: 27974160]

Sano K, Tsuda MC, Musatov S, Sakamoto T, and Ogawa S (2013). Differential effects of site-specific knockdown of estrogen receptor α in the medial amygdala, medial pre-optic area, and ventromedial nucleus of the hypothalamus on sexual and aggressive behavior of male mice. Eur. J. Neurosci 37, 1308–1319. [PubMed: 23347260]

Saper CB, and Lowell BB (2014). The hypothalamus. Curr. Biol. 24, R1111–R1116. [PubMed: 25465326]

Saper CB, Swanson LW, and Cowan WM (1976). The efferent connections of the ventromedial nucleus of the hypothalamus of the rat. J. Comp. Neurol 169, 409–442. [PubMed: 61975]

Satija R, Farrell JA, Gennert D, Schier AF, and Regev A (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol 33, 495–502. [PubMed: 25867923]

Saunders A, Macosko EZ, Wysoker A, Goldman M, Krienen FM, de Rivera H, Bien E, Baum M, Bortolin L, Wang S, et al. (2018). Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. Cell 174, 1015–1030 e1016. [PubMed: 30096299]

Seung HS, and Sumbul U (2014). Neuronal cell types and connectivity: lessons from the retina. Neuron 83, 1262–1272. [PubMed: 25233310]

Shah S, Lubeck E, Zhou W, and Cai L (2016). In Situ Transcription Profiling of Single Cells Reveals Spatial Organization of Cells in the Mouse Hippocampus. Neuron 92, 342–357. [PubMed: 27764670]

Shah S, Takei Y, Zhou W, Lubeck E, Yun J, Eng CL, Koulena N, Cronin C, Karp C, Liaw EJ, et al. (2018). Dynamics and Spatial Genomics of the Nascent Transcriptome by Intron seqFISH. Cell 174, 363–376 e316. [PubMed: 29887381]

Silva BA, Mattucci C, Krzywkowski P, Murana E, Illarionova A, Grinevich V, Canteras NS, Ragozzino D, and Gross CT (2013). Independent hypothalamic circuits for social and predator fear. Nat. Neurosci 16, 1731–1733. [PubMed: 24212674]

Simerly RB (2002). Wired for reproduction: organization and development of sexually dimorphic circuits in the mammalian forebrain. Annu. Rev. Neurosci 25, 507–536. [PubMed: 12052919]

Sommer C, Straehle C, Kothe U, and Hamprecht FA (2011). Ilastik: Interactive learning and segmentation toolkit. In 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 230–233.

Sternson SM (2013). Hypothalamic survival circuits: blueprints for purposive behaviors. Neuron 77, 810–824. [PubMed: 23473313]

Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, Hao Y, Stoeckius M, Smibert P, and Satija R (2019). Comprehensive Integration of Single-Cell Data. Cell 177, 1888–1902 e1821. [PubMed: 31178118]

Tasic B (2018). Single cell transcriptomics in neuroscience: cell classification and beyond. Curr. Opin. Neurobiol 50, 242–249. [PubMed: 29738987]

Tasic B, Menon V, Nguyen TN, Kim TK, Jarsky T, Yao Z, Levi B, Gray LT, Sorensen SA, Dolbeare T, et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat. Neurosci 19, 335–346. [PubMed: 26727548]

Tasic B, Yao Z, Graybuck LT, Smith KA, Nguyen TN, Bertagnolli D, Goldy J, Garren E, Economo MN, Viswanathan S, et al. (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature 563, 72–78. [PubMed: 30382198]

Tervo DG, Hwang BY, Viswanathan S, Gaj T, Lavzin M, Ritola KD, Lindo S, Michael S, Kuleshova E, Ojala D, et al. (2016). A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons. Neuron 92, 372–382. [PubMed: 27720486]

Tovote P, Esposito MS, Botta P, Chaudun F, Fadok JP, Markovic M, Wolff SB, Ramakrishnan C, Fenno L, Deisseroth K, et al. (2016). Midbrain circuits for defensive behaviour. Nature 534, 206–212. [PubMed: 27279213]

Tsuneoka Y, Tsukahara S, Yoshida S, Takase K, Oda S, Kuroda M, and Funato H (2017). Moxd1 Is a Marker for Sexual Dimorphismin the Medial Preoptic Area, Bed Nucleus of the Stria Terminalis and Medial Amygdala. Front. Neuroanat 11, 26. [PubMed: 28396628]

Wang L, Chen IZ, and Lin D (2015). Collateral pathways from the ventromedial hypothalamus mediate defensive behaviors. Neuron 85, 1344–1358. [PubMed: 25754823]

Wang L, Talwar V, Osakada T, Kuang A, Guo Z, Yamaguchi T, and Lin D (2019). Hypothalamic Control of Conspecific Self-Defense. Cell Rep 26, 1747–1758 e1745. [PubMed: 30759387]

Ward JH (1963). Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association 58, 236–244.

Wei YC, Wang SR, Jiao ZL, Zhang W, Lin JK, Li XY, Li SS, Zhang X, and Xu XH (2018). Medial preoptic area in mice is capable of mediating sexually dimorphic behaviors regardless of gender. Nat. Commun 9, 279. [PubMed: 29348568]

Welch JD, Kozareva V, Ferreira A, Vanderburg C, Martin C, and Macosko EZ (2019). Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity. Cell 177, 1873–1887 e1817. [PubMed: 31178122]

Wermter AK, Scherag A, Meyre D, Reichwald K, Durand E, Nguyen TT, Koberwitz K, Lichtner P, Meitinger T, Schäfer H, et al. (2008). Preferential reciprocal transfer of paternal/maternal DLK1 alleles to obese children: first evidence of polar overdominance in humans. Eur. J. Hum. Genet 16, 1126–1134. [PubMed: 18398438]

Wu Z, Autry AE, Bergan JF, Watabe-Uchida M, and Dulac CG (2014). Galanin neurons in the medial preoptic area govern parental behaviour. Nature 509, 325–330. [PubMed: 24828191]

Wu YE, Pan L, Zuo Y, Li X, and Hong W (2017). Detecting Activated Cell Populations Using Single-Cell RNA-Seq. Neuron 96, 313–329 e316. [PubMed: 29024657]

Xu Y, Nedungadi TP, Zhu L, Sobhani N, Irani BG, Davis KE, Zhang X, Zou F, Gent LM, Hahner LD, et al. (2011). Distinct hypothalamic neurons mediate estrogenic effects on energy homeostasis and reproduction. Cell Metab. 14, 453–465. [PubMed: 21982706]

Xu X, Coats JK, Yang CF, Wang A, Ahmed OM, Alvarado M, Izumi T, and Shah NM (2012). Modular genetic control of sexually dimorphic behaviors. Cell 148, 596–607. [PubMed: 22304924]

Yang CF, and Shah NM (2014). Representing sex in the brain, one module at a time. Neuron 82, 261–278. [PubMed: 24742456]

Yang CF, Chiang MC, Gray DC, Prabhakaran M, Alvarado M, Juntti SA, Unger EK, Wells JA, and Shah NM (2013). Sexually dimorphic neurons in the ventromedial hypothalamus govern mating in both sexes and aggression in males. Cell 153, 896–909. [PubMed: 23663785]

Yang T, Yang CF, Chizari MD, Maheswaranathan N, Burke KJ Jr., Borius M, Inoue S, Chiang MC, Bender KJ, Ganguli S, et al. (2017). Social Control of Hypothalamus-Mediated Male Aggression. Neuron 95, 955–970 e954. [PubMed: 28757304]

Yu JY, Kanai MI, Demir E, Jefferis GSXE, and Dickson BJ (2010). Cellular organization of the neural circuit that drives Drosophila courtship behavior. Curr. Biol 20, 1602–1614. [PubMed: 20832315]

Zeisel A, Hochgerner H, Lonnerberg P, Johnsson A, Memic F, van der Zwan J, Haring M, Braun E, Borm LE, La Manno G, et al. (2018). Molecular Architecture of the Mouse Nervous System. Cell 174, 999–1014 e1022. [PubMed: 30096314]

Zeng H, and Sanes JR (2017). Neuronal cell-type classification: challenges, opportunities and the path forward. Nat. Rev. Neurosci 18, 530–546. [PubMed: 28775344]

Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun 8, 14049. [PubMed: 28091601]

Highlights

- SMART-seq and 10x platforms reveal 17 transcriptomic cell types (T-types) in VMHvl
- Multiplexed smFISH confirms scRNA-seq and identifies T-types with A-P restrictions
- Rare female- and male-specific T-types are found among the $Esr1^{+}$ subclusters
- Relatively few VMHvl T-types exhibit behavior-specific activation and connectivity

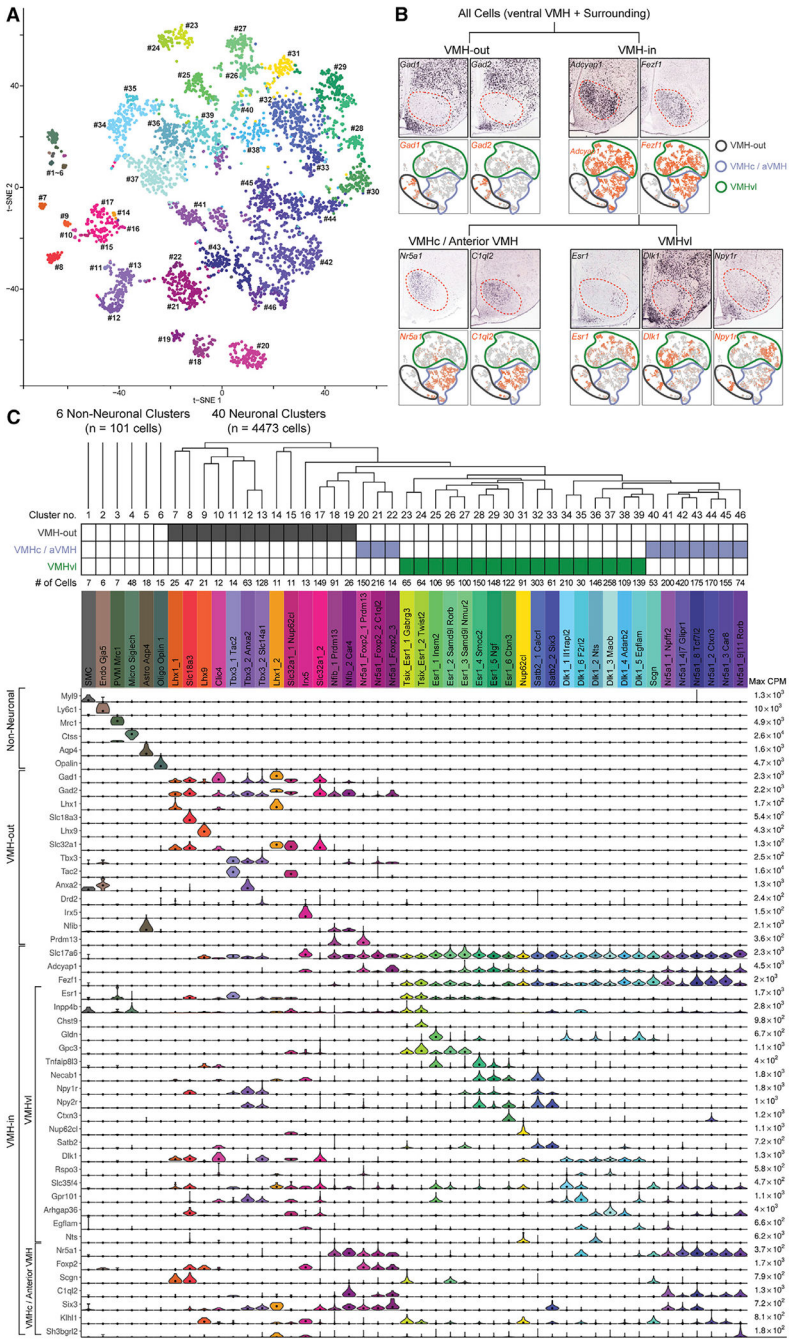


Figure 1. VMHvl Transcriptomic Cell Types Revealed by SMART-Seq

(A) Two-dimensional t-distributed stochastic neighbor embedding (t-SNE) plot color-coded by 46 SMART-seq clusters (N = 4,574) from ventral VMH.

(B) Expression patterns of major marker genes in VMH (red dashed circle) and surrounding areas by ISH (Allen Mouse Brain Atlas; upper) arranged by anatomical hierarchicy; orange points on t-SNE plots (lower) indicate their expression levels. Anatomic location of clusters is outlined on t-SNE plots (VMH-out: black; VMHvl: green; VMHc or anterior VMH: light blue).

(C) Violin plots illustrating expression levels of marker genes by cluster; “max CPM” (right), maximum counts per million reads. Dendrogram and matrix show relatedness between clusters and their spatial locations (see B), respectively.
See also Figures S1 and S2.

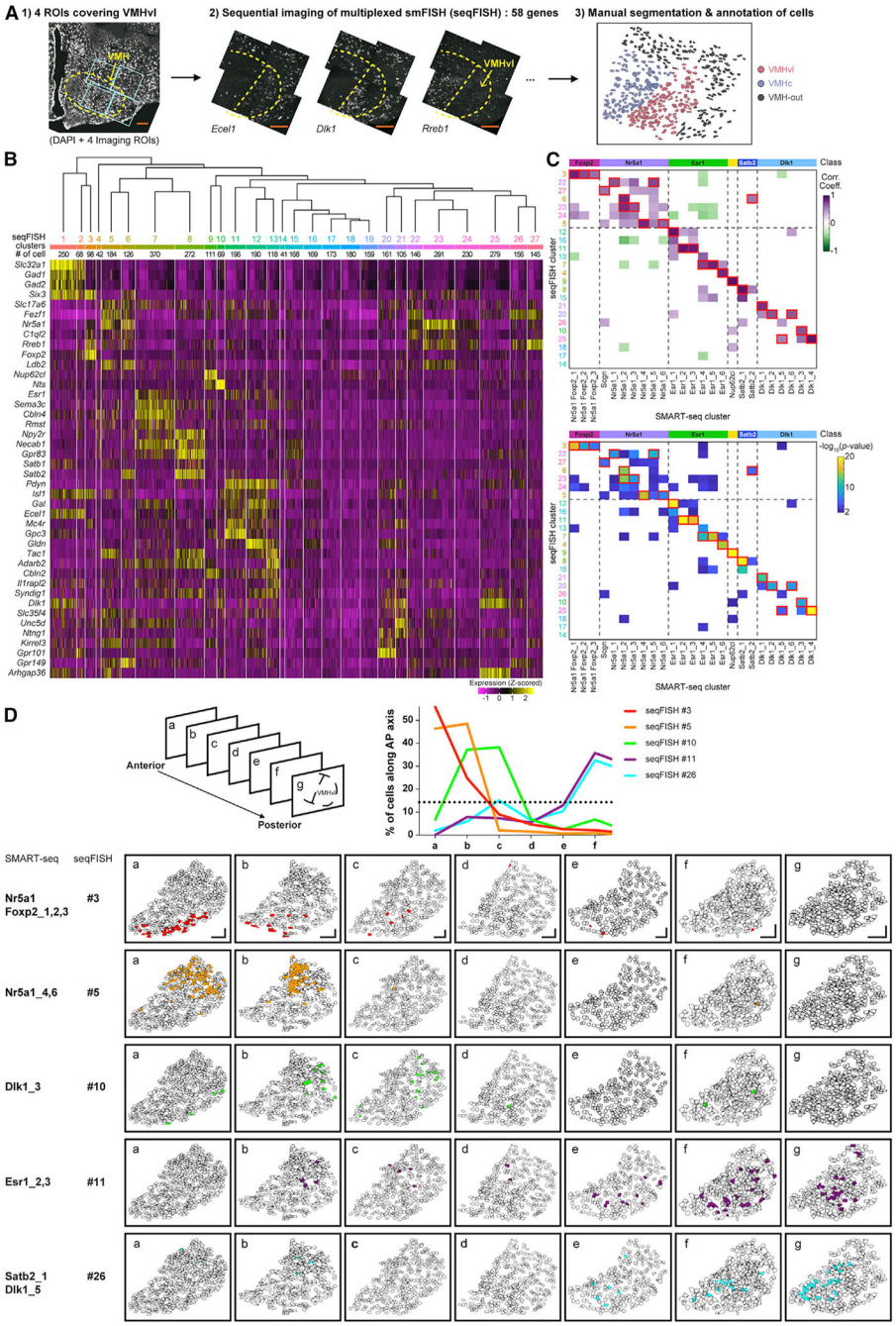


Figure 2. Anatomic Distribution of VMH scRNA-Seq Clusters by seqFISH

(A) Schematic of seqFISH procedure in VMH. Light blue solid lines (1) outline ROIs of sequential hybridizations; yellow dashed lines (2) (maximum intensity Z projections) outline VMH and VMHvl. Three major anatomic regions (VMHvl, VMHc, and VMH-out) are color coded (3). Scale bars, 100 μm (inset).

(B) Heatmap showing expression level of marker genes (rows) in 27 seqFISH clusters (columns; n = 4,497; VMHvl only).

(C) Heatmaps showing correlation between seqFISH and SMART-seq clusters (n = 3,824; VMH only) and their p values (white, $p > 0.05$ or $r < 0.35$; see STAR Methods). seqFISH clusters that are the most strongly correlated with each scRNA-seq cluster are marked by red squares.

(D) Spatial distribution along A-P axis of seqFISH clusters (color-filled) showing anterior (#3, #5, #10) or posterior (#11, #26) biases, projected onto all segmented VMHvl cells (a–f). Line plot shows quantification for indicated clusters; black dotted line shows chance distribution. Scale bars, 50 μm (inset).

See also Figure S3.

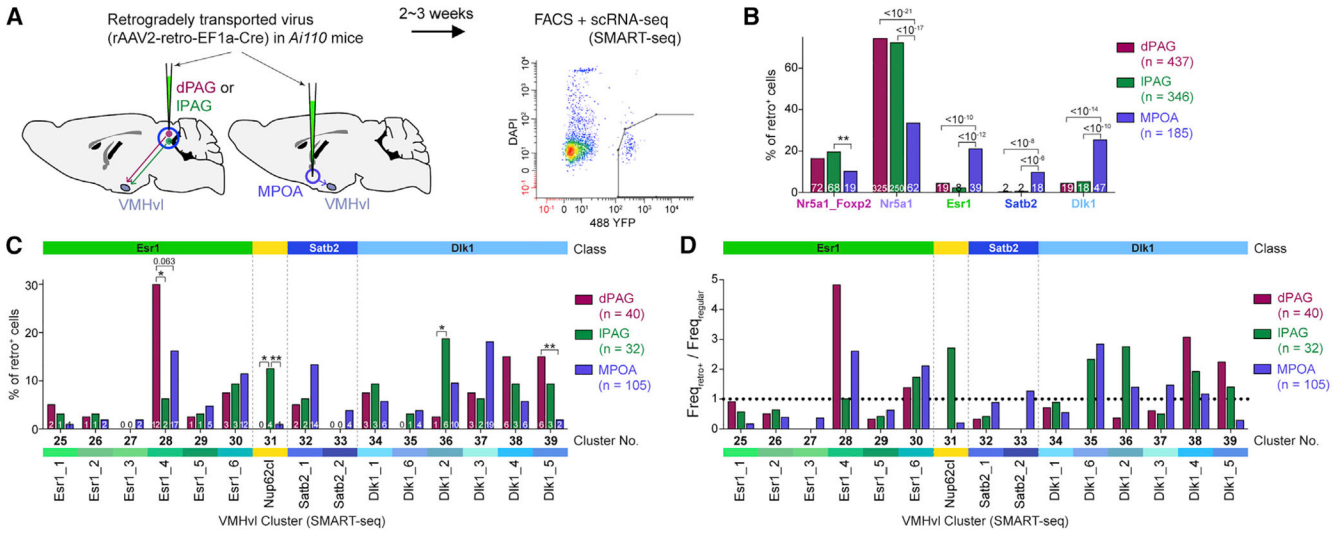


Figure 3. Anterior versus Posterior Projection Biases of Neurons in scRNA-seq Clusters

(A) Schematic of Retro-seq procedure (see STAR Methods) from dPAG, IPAG, or MPOA.

(B) Bar plot shows the distributions of cells retrogradely labeled from each target for the major VMH classes.

(C and D) Percent of retrogradely labeled cells (C) and their relative frequency compared to non-Retro-seq samples (D) in each VMHvl SMART-seq cluster. Numbers of retro^{+} cells sequenced are listed at the bottom of the graphs.

*p < 0.05, **p < 0.01 (Fisher's exact test).

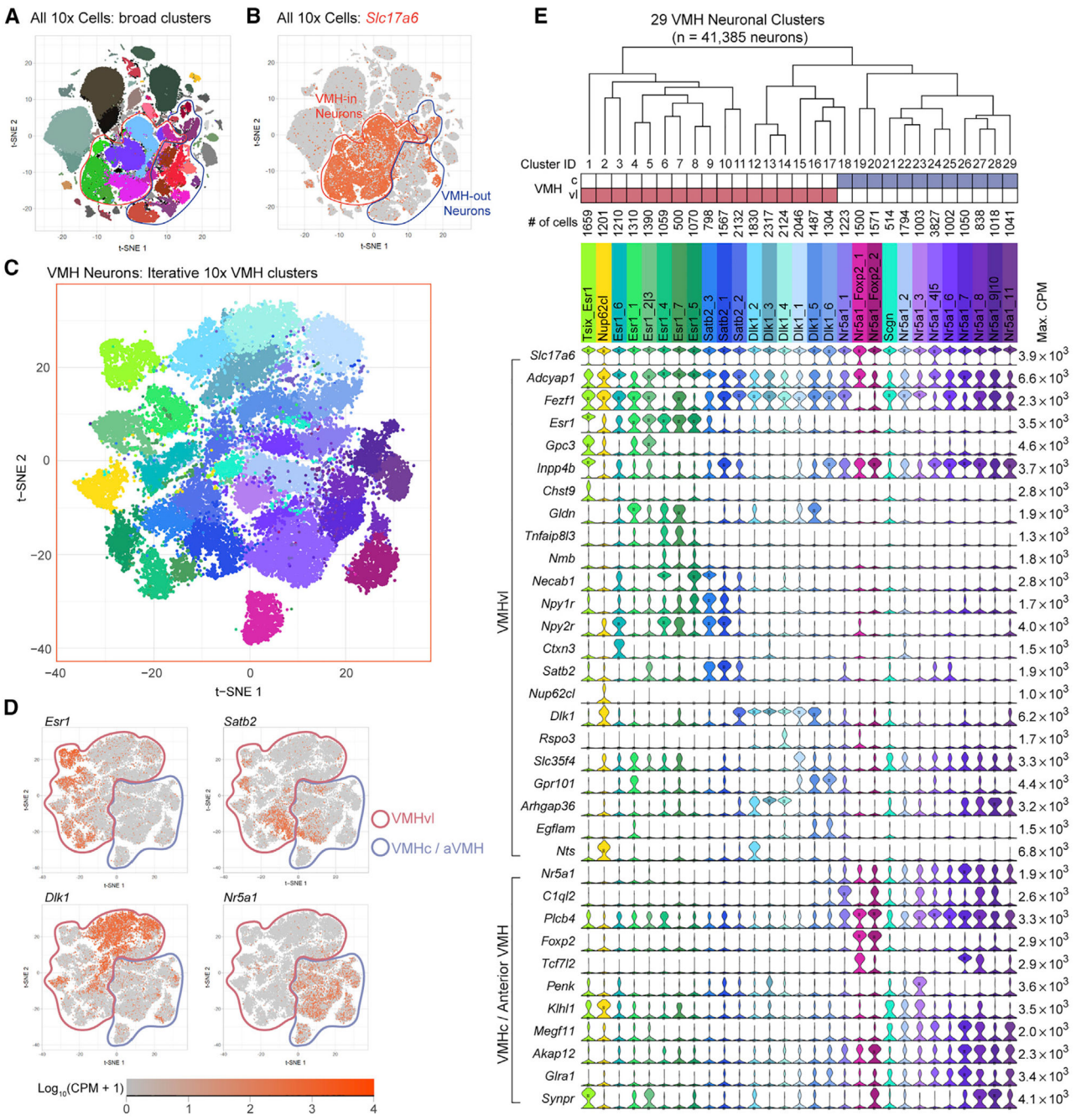


Figure 4. VMHvl Transcriptomic Cell Types Revealed by 10x scRNA-seq

(A and B) t-SNE plots showing clusters (A) and their *Vglut2* (*Slc17a6*) expressions (B) for all cells analyzed (N = 149,663). All neuronal clusters (n = 78,476) are outlined in either red (VMH-in) or blue (VMH-out), respectively.

(C) t-SNE plot illustrating results of iterative clustering of VMH-in cells (n = 41,385; 29 clusters).

(D) Expression levels of four main class-specific marker genes (*Esr1*, *Satb2*, *Dlk1*, and *Nr5a1*) are color coded (orange) on t-SNE plots; VMH subdivisions are outlined in different colors (VMHvl: light red; VMHc or anterior VMH: light blue).

(E) Violin plots show differential expressions of marker genes (VMHvl versus VMHc-enriched; indicated by left brackets) among 29 VMH clusters (C) with their spatial locations (top matrix).

See also Figure S4.

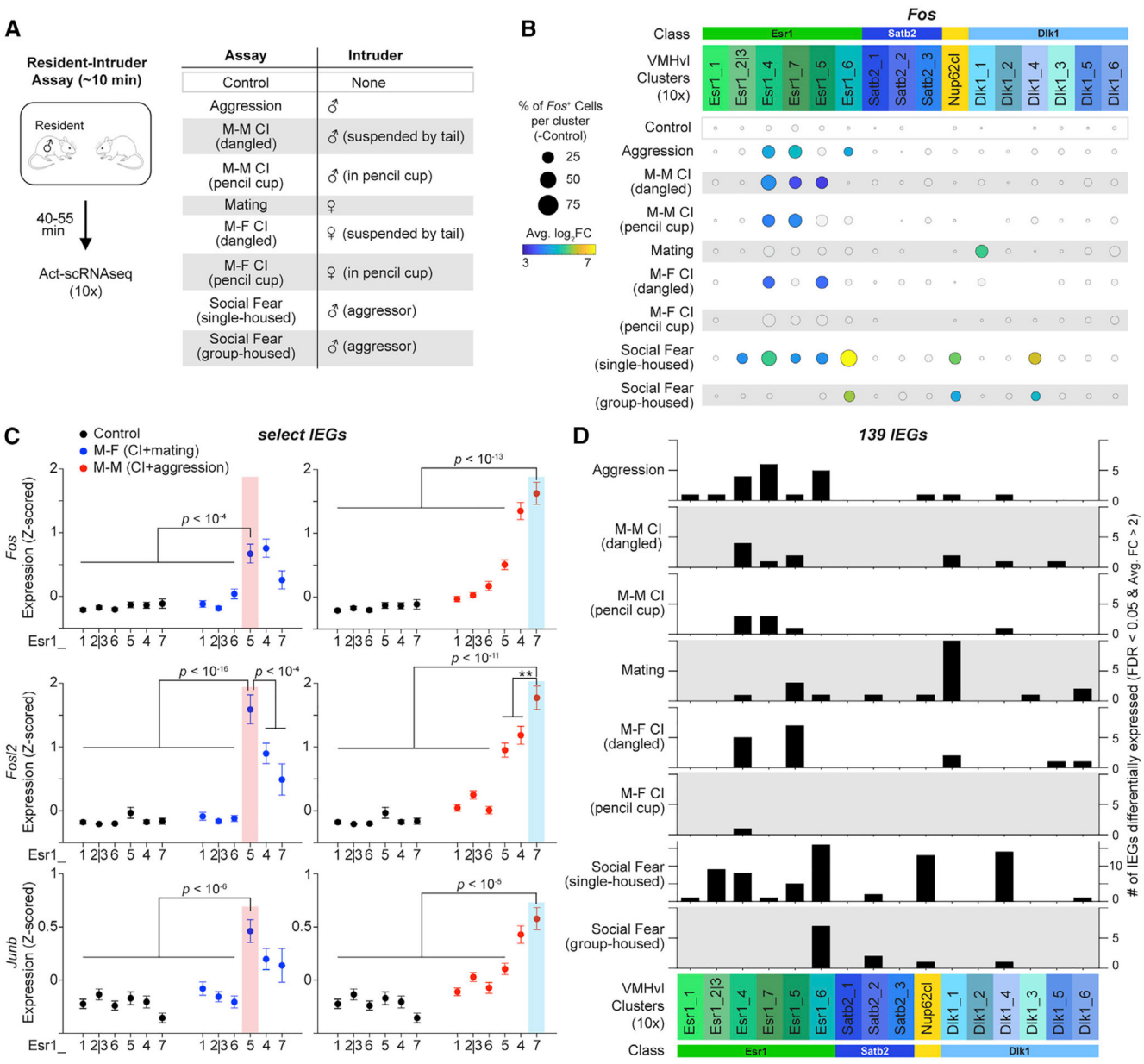


Figure 5. Activation of Neurons in VMHvl Clusters during Different Social Behaviors

(A) Schematic of Act-seq protocol and summary of behavioral assays (see STAR Methods; Figure S4I).

(B) Dot plots illustrating *Fos* induction in 10x VMHvl clusters. Colored and shaded gray dots indicate clusters with a significant ($p < 0.05$ after multiple comparison corrections across behaviors and clusters; three-way ANOVA and Bonferroni post hoc test), and non-significant induction ($p > 0.05$) versus control (or clusters from control animals), respectively. Dot size indicates proportion of *Fos*$^+$ cells per cluster after subtraction of control values; color indicates average $\log_2$fold-change (FC) versus control.

(C) Expression levels of *Fos*, *Fosl2*, and *Junb* in 6 Esr1 clusters (male only). The clusters showing the highest induction of *Fos*, *Fosl2*, or *Junb* by either M-M (aggression plus M-M

CI tests; red points) or M-F (mating plus M-F CI tests; blue points) versus control (black points) are highlighted by red (left; Esr1_5) or blue (right; Esr1_7) shading, respectively (**p < 0.01, ***p < 0.001; three-way ANOVA and Bonferroni post hoc test; data are represented as mean ± SEM).

(D) Bar graphs showing the number of different IEGs significantly induced during each social behavior in each cluster (see also Figure S5B) compared with control animals, where average FC > 2 and false discovery rate (FDR) <0.05.

See also Figures S5 and S6.

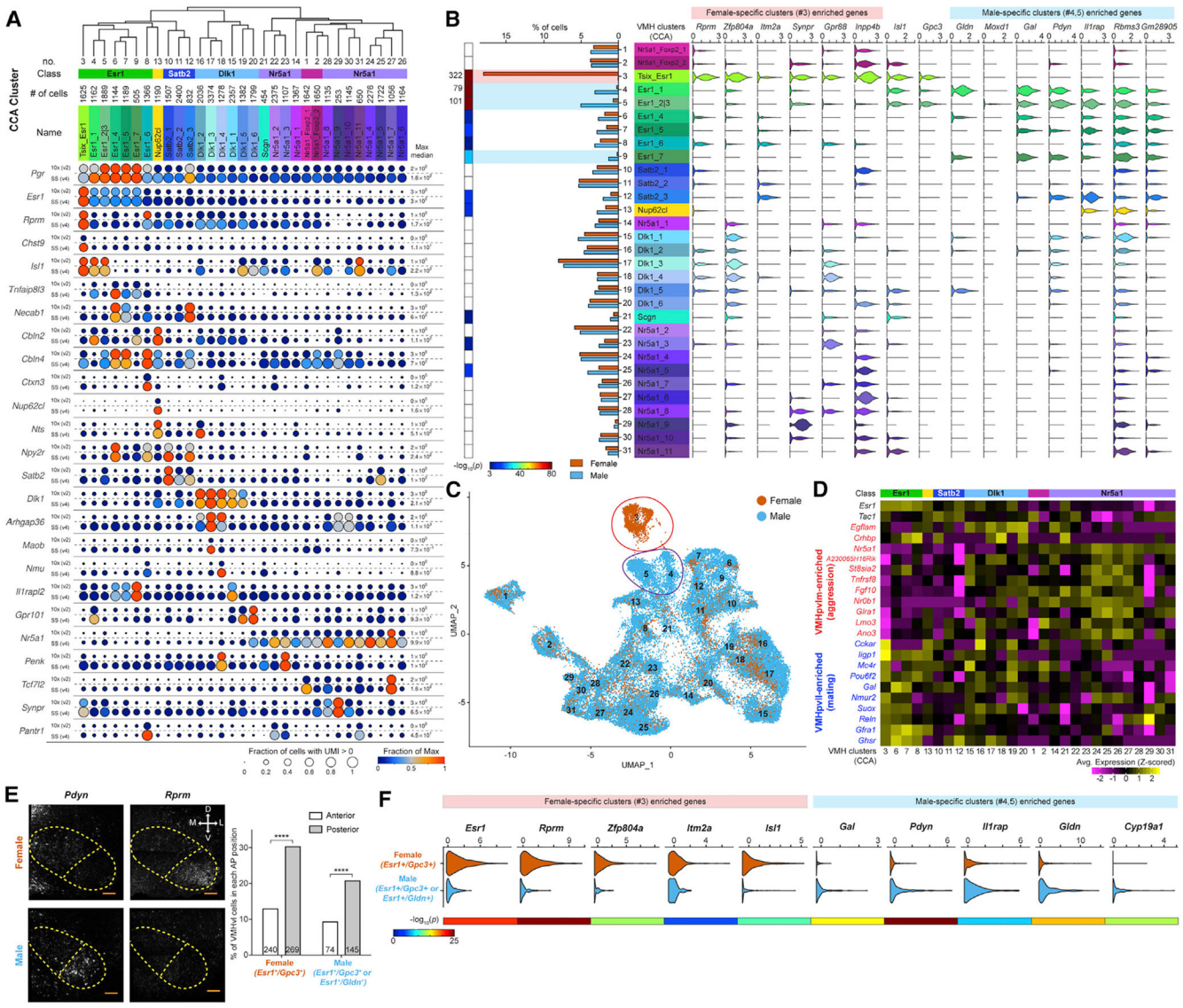


Figure 6. CCA Clusters and Sexually Dimorphic Cell Types in VMHvl

(A) Dot plots illustrating marker gene expressions in joint CCA clusters for SMART-seq (bottom) and 10x (top) datasets. Dot size and color indicate proportion of expressing cells and average expression level in each cluster, respectively.

(B) Bar graph (left) represents the fractions of cells in clusters by sex (orange-red, female; sky blue, male). Far left, log-scaled p values ($-\log_{10}p$) are color coded (Fisher's exact test; white, p > 0.001). Violin plots show expression of genes specifically enriched in either female-specific (#3) or male-specific (#4, #5, #9) *Esr1* clusters.

(C) t-SNE plots of VMH cells from males (sky blue) and females (orange-red) illustrating distribution across joint CCA clusters.

(D) Heatmap showing average expression in female VMH CCA clusters (n = 8,793) of genes previously identified by bulk RNA-seq (Hashikawa et al., 2017a) as localized to aggression-activated (red) or mating-activated (blue) VMHvl subdivisions in females (VMHpvlm versus VMHpvll).

(E) Examples of seqFISH images (left; maximum intensity Z projections) for *Pdyn* and *Rprm*, major marker genes differentially expressed between male and female VMHvl. Scale bars, 100 μm (inset). Bar graph (right) showing spatial distribution patterns (along A-P axis) of female *Esr1*$^+$ and *Gpc3*$^+$ or male *Esr1*$^+$ and *Gpc3*$^+$ or *Esr1*$^+$ and *Gldn*$^+$ cells in VMHvl (****$p < 0.0001$; Fisher's exact test).

(F) Violin plots showing differential expressions of 10 genes (previously identified by scRNA-seq) between female- (*Esr1*$^+$/*Gpc3*$^+$) and male- (*Esr1*$^+$/*Gpc3*$^+$ or *Esr1*$^+$/*Gldn*$^+$) specific cell types in seqFISH. Log-scaled p values ($-\log_{10}p$) are color coded (below; unpaired t test).

See also Figure S7.

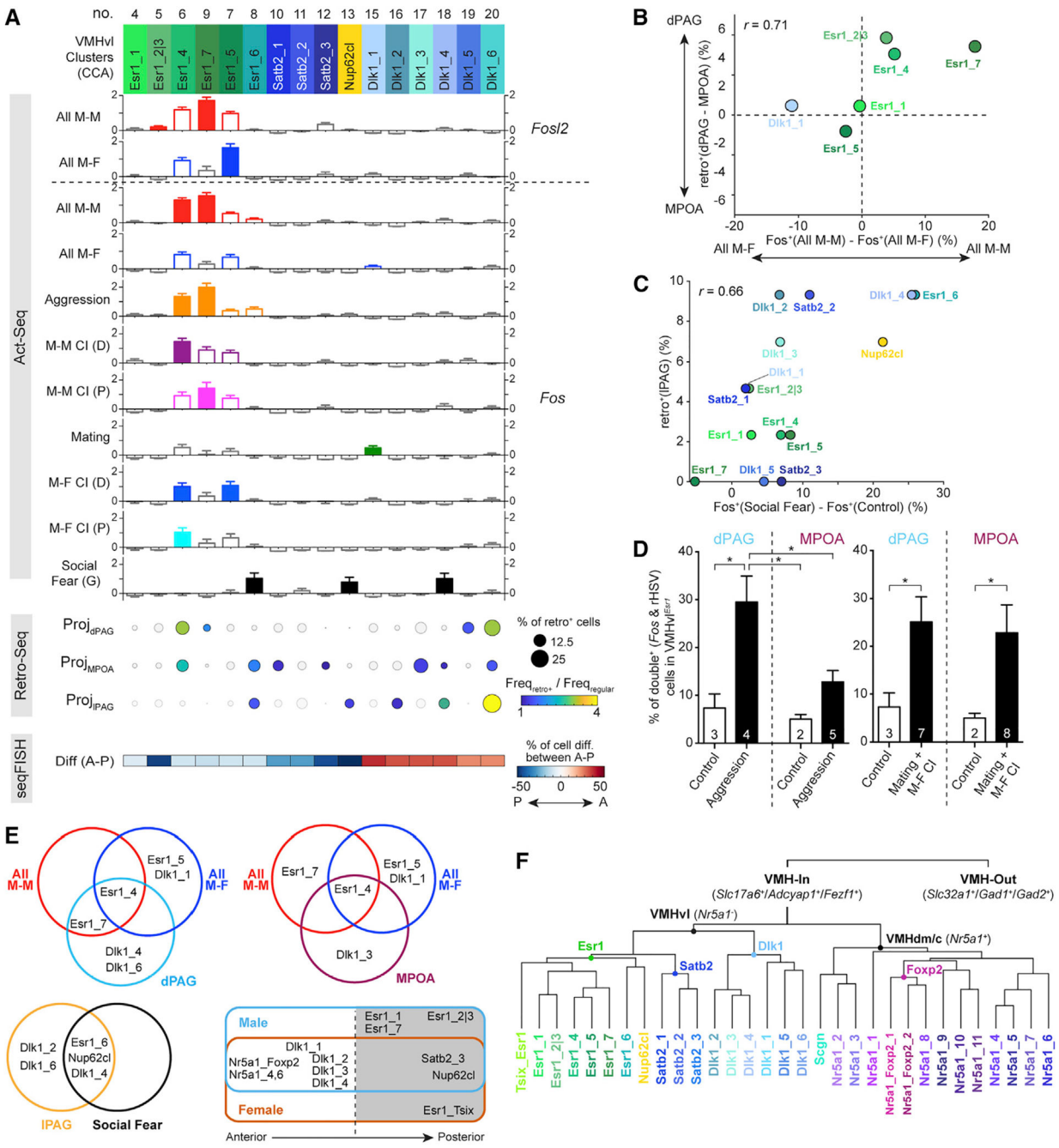


Figure 7. Relationship between Behavioral Activation, Projection Specificity, and Transcriptomic Identity in the CCA Framework

(A) Bar graphs, dot plots, and heatmap illustrating Act-seq (top; Figure 5B), Retro-seq (middle; Figures 3C and 3D), and seqFISH (bottom; Figure 2D) data projected onto VMHvl CCA clusters, respectively. Act-seq: bars with colored and gray outlines indicate significant ($p < 0.05$ after multiple-comparison corrections across behaviors and clusters; three-way ANOVA and Bonferroni post hoc test), and non-significant ($p > 0.05$) differences in IEG expression versus control, respectively. Filled and open bars indicate significant and non-significant differences in IEG expression versus other clusters within a given behavior or

behavioral category, respectively. All M-M and all M-F indicate combined data from all male-male or male-female social behaviors, respectively, for both *Fos* and *Fosl2* expression. Lower bar graphs illustrate data for individual behaviors (*Fos* only). Retro-seq: dot size indicates the percent of retrogradely labeled VMH cells from dPAG, lPAG, or MPOA (middle); dot colors indicate relative ratio (retrogradely labeled divided by total sequenced populations) of clusters. Shaded gray dots indicate clusters with ratio <1). seqFISH: "Diff (A-P)" indicates percent difference in cell number between anterior versus posterior VMHvl (color scale) for cells in corresponding seqFISH clusters.

(B and C) Scatterplots showing correlation between preferential *Fos* activation during all M-M versus all M-F behaviors (x axis) and projection preference between dPAG versus MPOA (y axis) for major Act-seq clusters (B) and correlation between preferential *Fos* activation during social fear (using group-housed mice) versus control (x axis) and proportion of cells projecting to lPAG (y axis) for VMHvl CCA clusters (C).

(D) Bar graph showing the proportion of double^{+} (*Fos*$^{+}$ and *rHSV*$^{+}$) cells among VMHvl^{Esr1} neurons between aggression versus control (upper) or M-F social interaction tests (mating + M-F CI tests) versus control (lower) in *Esr1*-Cre mice injected in dPAG or MPOA with a Cre-dependent retrograde HSV (rHSV). Numbers of mice are listed at the bottom of the graphs. *$p < 0.05$, one-way ANOVA after multiple-comparison corrections (left) and unpaired t test (right).

(E) Venn diagrams (1–3) illustrating joint CCA clusters activated during different behavioral categories (all M-M, all M-F, and social fear) or projecting to dPAG and MPOA versus lPAG, respectively. Venn diagram (4) illustrates clusters distinctively located along A-P axis in male and female VMHvl.

(F) Summary hierarchical taxonomy of CCA T-types in VMH; cell classes are labeled at branch points of the dendrogram.

Data in (A) and (D) are represented as mean ± SEM. CI, close investigation (see Figure 5).

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Goat polyclonal anti-*c-fos*	Santa Cruz Biotechnology	Cat# sc-52-g; RRID: AB_2629503
Donkey anti-goat, Alexa Fluor 647	Invitrogen	Cat# A-21447; RRID: AB_2535864
Bacterial and Virus Strains		
rAAV2-retro-EF1a-Cre	Tervo et al., 2016	N/A
HSV1-LS1L-EYFP	MGH Vector Core	N/A
Chemicals, Peptides, and Recombinant Proteins		
Pronase	Sigma	Cat# P6911-1G
Papain	Sigma	Cat# P3125-250MG
Egg White/BSA ovomucoid inhibitor	Worthington	Cat# OI-BSA
DNase I	Thermo scientific	Cat# 90083
DNase/RNase-Free Distilled Water	Thermo scientific	Cat# 10977-015
10X PBS	Thermo scientific	Cat# AM9624
DAPI	Sigma	Cat# D9542-10MG
BSA	Thermo scientific	Cat# AM2616
Actinomycin D	Sigma	Cat# A1410-2MG
Kynurenic acid sodium salt	Abcam	Cat# ab120256
L-Cysteine	Sigma	Cat# W326305-100G
D-(+)-Trehalose dihydrate	Sigma	Cat# 90210-50G
Ethylenediaminetetraacetic acid	Sigma	Cat# 03690-100ML
$MgSO_4$solution	Sigma	Cat# 83266-100ML-F
$CaCl_2$solution	Sigma	Cat# 21115-100ML
Sodium nitrite	Sigma	Cat# 237213-100G
Heparin sodium salt	Sigma	Cat# H4784-250MG
Paraformaldehyde	EMS	Cat# 15714-S
Sucrose, RNase & DNase Free	Amresco	Cat# 0335-2.5KG
Critical Commercial Assays		
SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing	Takara	Cat# 634894
Nextera XT Index Kit V2 Set A	Illumina	FC-131-2001
Nextera XT DNA Library Preparation	Illumina	FC-131-1096
Chromium Single Cell 3′ Library & Gel Bead Kit v2	10x Genomics	PN-120237
Chromium Single Cell 3′ Chip Kit v2	10x Genomics	PN-120236
Chromium i7 Multiplex Kit	10x Genomics	PN-120262
Experimental Models: Organisms/Strains		
Mouse: *Nr5a1-Cre*	Dhillon et al., 2006	N/A
Mouse: *Oxtr-T2A-Cre*	Daigle et al., 2018	N/A
Mouse: *Vglut2-ires-cre*	The Jackson Laboratory	JAX: 016963

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse: *Esr1-Cre*	Lee et al., 2014	N/A
Mouse: *Ai14*(RCL-tdTomato)	Madisen et al., 2010	N/A
Mouse: *Ai110*(RCL-FnGF-nT)	Daigle et al., 2018	N/A
Mouse: C57BL/6N	Charles River	N/A
Mouse: BALB/c	Charles River	N/A
Software and Algorithms		
CellRanger v3.0.2	10x Genomics	http://www.10xgenomics.com/
MATLAB Version 2017a	MathWorks	https://www.mathworks.com
R v.3.3.0	R Foundation	https://www.r-project.org
Seurat v3.0.3	Butler et al., 2018; Stuart et al., 2019	https://satijalab.org/seurat/
scrattch.hicat	Tasic et al., 2018	https://github.com/AllenInstitute/scrattch.hicat
seqFISH image processing code (MATLAB)	Shah et al., 2016; Eng et al., 2019	N/A
SMART-seq, 10x, and seqFISH data analysis code code (MATLAB and R)	This paper	N/A
Prism 6	GraphPad Software	https://www.graphpad.com/
ImageJ	NIH	https://imagej.nih.gov/ij